一种基于多变换的LSB隐写算法*)

徐江峰[1] 李 昊[1] 杨 有[2]

(郑州大学信息工程学院 郑州 450001)[1] (重庆师范大学数学与计算机科学学院 重庆 400047)[2]

摘 要 本文对传统的LSB隐写算法进行了研究,提出了一种新的LSB隐写方法,该方法利用多种变换实现信息隐藏。实验及理论分析结果表明,该方法可以抵御RS分析及统计分析的攻击,提高传统LSB隐写算法的安全性。

关键词 信息隐藏,LSB算法,密写分析,安全性

An Improve LSB Hide Algorithm Based on Multi-Transformation

XU Jiang-Feng[1] LI Hao[1] YANG You[2]

(School of Information and Engineering, Zhengzhou University, Zhengzhou 450001)[1]

(College of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047)[2]

Abstract This paper firstly introduces the traditional LSB information hiding scheme, then proposes an improved LSB algorithm. This algorithm implements the information by many kinds of transformation. Theoretical analysis and experimental results demonstrate that the information hiding scheme can effectively resist both RS and statistics analyses, and improve the security of traditional LSB scheme.

Keywords Information hiding, LSB algorithm, Steganalysis, Security

1 引言

随着网络技术的快速发展,信息安全问题变得越来越突出。为了保护秘密数据,数据加密技术得到了广泛应用,然而加密后的密文常常是一些不可解读的乱码,使得在保护数据的同时也易引起攻击者的注意。为了解决这一问题,信息隐藏技术得到了研究和发展,并且取得了许多研究成果[1~5]。隐写是重要的隐藏技术之一,它把秘密信息隐藏在含有大量冗余信息的载体图像中,但从视觉效果来看,隐写图像和载体图像并没有区别。基于图像的信息隐藏算法分为空域和频域两大类,其中LSB算法和DCT算法分别是它们的代表。一个安全的隐写算法仅限于视觉不可见性是不够的,攻击者会通过分析隐写图像和载体图像统计特性的变化来分析图像中是否隐含有隐密信息。

传统的LSB算法变换简单,但它不能抵御RS分析技术[6,7]和统计分析技术的攻击[8]。本文在对传统LSB算法特性进行分析的基础上,提出了一个新的LSB隐写算法,并对算法的安全性进行了理论和实验分析。分析结果表明,该算法提高了传统LSB算法的安全性,可以抵御基于RS分析和统计分析技术的攻击。

2 传统LSB隐写算法

传统LSB密写方案利用载体中的最不重要位来隐藏秘密信息,它利用人眼的不敏感性,将隐秘信息隐藏在图像之后。对于256级灰度图像$G=(g_{ij})_{M\times N}$,用横竖间隔相等的两组直线把图像分割成大小相等的像素块,设M'、N'分别表示图像划分后每行的块数和每列的块数,则$G=(B_{ij})_{M'\times N'}$,其中B_{ij}表示划分后图像的第i行、第j列对应的像素块。若要隐藏长度为s的二进制信息,则需从$M'\times N'$中选取s个图像块,并从每个块中取出一个像素点,利用像素的最低位隐藏二进制信息。

根据隐藏时隐秘信息隐藏位置选择方法的不同,LSB隐写算法可分为顺序嵌入、随机嵌入和矩阵编码法嵌入等。顺序嵌入是一种最为简单的方法,但其安全性较低,因为利用该方法得到的隐写图像,嵌入信息的部分与没有嵌入信息的部分统计特征有明显的差异,有许多掩密分析方法都可以检测到信息隐藏的事实。为了抵抗这些攻击,顺序嵌入被随机嵌入和矩阵编码嵌入替代。

利用随机的方式将隐秘信息嵌入到载体中时,信息比较均匀地分散到了载体中,这样就无法区分载体中的嵌入信息部分和未嵌入信息部分。而采用矩阵编码法则减少了最不重要位改变的数量。这两种方法均提高了掩密的安全性。然而,即使如此,也有不少掩密分析方法可以通过被动攻击或主动攻击检测到信息的存在[9,10]。

3 改进的LSB算法

要实现信息的隐藏和提取,通常采用如下步骤:

(1)对隐秘信息M进行加密预处理,得到$M'=E(K_1,M)$,其中E为加密算法,K_1为加密密钥。

(2)利用嵌入算法P把信息M'嵌入到载体对象C中,得到隐藏对象$C'=P(K_2,M',C)$,其中K_2是隐藏密钥。

(3)发送者A把隐藏对象C'传递给接收者B。

(4)接收者B对接收到的隐藏对象C',利用提取算法Q及密钥K_2得到隐秘信息$M'=Q(K_2,C',C)$或$M'=Q(K_2,C')$;

(5)接收者B利用解密算法D及密钥K_1得到真正的隐秘信息$M=D(K_1,M')$。

从上述步骤中可以看出,隐藏信息的安全性主要取决于

*)本课题得到河南省教委自然科学基金支持(2006520014)。徐江峰 博士,副教授,主要研究方向为数据加密技术;李 昊 硕士,研究方向为信息隐藏技术;杨 有 博士研究生。

密钥 K_1、K_2、加密算法 E 及隐藏算法 P。其中密钥 K_1 及加密算法 E 的安全性属于密码学研究范畴，而 K_2 主要确定隐藏信息在载体中的位置。

设 F_1 表示 $2i\leftrightarrow 2i+1$ 变换，F_{-1} 表示 $2i\leftrightarrow 2i-1$ 变换，$G=(g_{ij})_{M\times N}$ 是一个 256 级的灰度图像，$M=m_1,m_2,\cdots,m_n$ 是要隐藏的二进制信息，则新的 LSB 信息隐藏算法可以描述为：

(1)输入载体图像 $G=(g_{ij})_{M\times N}$，令 $C=G$；

(2)输入要隐藏的二进制信息 $M=m_1m_2,\cdots,m_n$；

(3)若 n 为奇数，则令 $n=n+1$，$M=m_1m_2,\cdots,m_nm_{n+1}$，其中 $m_{n+1}=0$；

(4)确定 G 中信息隐藏的位置 $K_2=(p_1,p_2,\cdots,p_{n/2})$；

(5)$i=1$，$flag=0$；

(6)若 $i\leqslant n/2$，则

i)取出 C 中位置 p_i 对应的灰度值 $c=a_7a_6\cdots a_1a_0$，$0\leqslant a_i\leqslant 1(i=0,1,\cdots,7)$；

ii)计算 $t=(m_{2i-1}m_{2i})\oplus(a_1a_0)$

iii)如果 $t=00$，则 $c'=c$；

如果 $t=01$，则 $c'=F_1(c)$；

如果 $t=10$，则

当 $flag=0$ 时，$c'=F_1(F_{-1}(c))$，$flag=1$；

当 $flag=1$ 时，$c'=F_{-1}(F_1(c))$，$flag=0$；

如果 $t=11$，则 $c'=F_{-1}(c)$；

iv)用 c' 代替 C 中位置 p_i 对应的灰度值 c；

v)$i=i+1$；

vi)返回(6)。

(7)输出隐秘图像 C

当隐藏图像为图 1，隐藏信息 $M=$“郑州大学信息工程学院”时，利用上述算法得到的隐秘图像如图 2 所示，其中隐藏位置 $K_2=(p_1,p_2,\cdots,p_{n/2})$ 是随机选取的。

若要提取隐藏图像中的信息，只要知道了隐藏信息的位置及顺序，取出对应点灰度值的最后两位，组成二进制序列即可。

图 1　原始图像

图 2　隐秘图像

4　性能分析

对于传统的 LSB 隐写，已有多种分析技术可以对其进行有效攻击，不但能判断出图像是否进行过 LSB 隐写，而且还可以计算出隐藏信息量。下面我们首先介绍两种对 LSB 隐写有效的攻击方法，而后分析这些技术对改进算法攻击的有效性。

4.1　RS 掩密分析技术

RS 分析方法是由 Jessica Fridrich 等人提出的[6,7]，是一种针对 LSB 隐写的分析方法，它不但能判断出图像中是否隐藏有隐秘信息，还能比较精确地估计出嵌入信息的长度。

对于给定的图像，函数 $f(x)=\sum|x-x_1|+\sum|x-x_2|$ 表示图像的混乱程度，其中 x 是图像块的灰度矩阵，x_1、x_2 分别是将 x 循环左移一列和循环下移一行得到的。从定义可以看出，$f(x)$ 的值越大，表明该图像越混乱，相邻像素间的相关性越小。

在进行 RS 分析时，首先把得到的图像分为大小相等的若干块，并从每个图像块中随机抽取部分像素进行 F_1 变换，而后计算变换后 $f(x)$ 的值，分别求出混乱度增加和减小图像块所占比例 R_+、S_+。同样地，进行 F_{-1} 变换后再计算出混乱度增加和减小的图像块的所占比例 R_- 和 S_-。

如果图像没有经过 LSB 隐写，那么无论用 F_1 变换还是 F_{-1} 变换，混乱度的改变基本相同，即满足 $R_+\approx R_-$ 和 $S_+\approx S_-$。而对隐藏了隐秘信息的图像来说，上式是不成立的，并且随着信息嵌入率的不断增大 R_+ 与 R_-、S_+ 与 S_- 的差别越来越大。

4.2　统计分析法

该分析方法是由邓倩岚和林家骏提出的[8]。设秘密信息的负载率为 α，即载体图像每个像素上平均负载 α 比特隐秘信息，原始图像中灰度值为 j 的像素数目为 f_j，则根据 LSB 隐写的原理，大约 $\alpha f_{2i}/2$ 个像素的灰度值由 $2i$ 改为 $2i+1$，也有约 $\alpha f_{2i+1}/2$ 个像素的灰度值由 $2i+1$ 改为 $2i$。由此得到

$$E(h_{2i})=f_{2i}-\alpha f_{2i}/2+\alpha f_{2i+1}/2$$

$$E(h_{2i+1})=f_{2i+1}-\alpha f_{2i+1}/2+\alpha f_{2i}/2$$

$$E(h_{2i+2})=f_{2i+2}-\alpha f_{2i+2}/2+\alpha f_{2i+3}/2$$

其中 E() 表示数学期望，从而有

$$E(h_{2i+1}-h_{2i})=(1-\alpha)(f_{2i+1}-f_{2i})$$

$$E(h_{2i+2}-h_{2i+1})=(f_{2i+2}-f_{2i+1})+\alpha(f_{2i+3}-f_{2i+2}+f_{2i+1}-f_{2i})/2$$

由此可知，随着嵌入率 α 的增大，h_{2i+1} 与 h_{2i} 趋近相等，而 h_{2i+2} 与 f_{2i+1} 则不是如此。因此，在得到一幅待检图像后，首先计算 $H_1=\sum_i|h_{2i+1}-h_{2i}|$ 和 $H_2=\sum_i|h_{2i+2}-h_{2i+1}|$ 的值，而后再对该图像嵌入测试信息，并计算嵌入信息后 H_1、H_2 的值 H'_1、H'_2，通过比较 $R=\frac{H_2}{H_1}$ 和 $R'=\frac{H'_2}{H'_1}$ 的大小即可知道待检图像是否经过 LSB 隐写，隐秘信息量有多大。

4.3　新算法安全性分析

(1)抵御 RS 分析

在传统的 LSB 隐写算法中，一般进行的只有 F_1 或 F_{-1} 变换，而上述算法中进行的有 F_1、F_{-1}、F_1F_{-1}、$F_{-1}F$ 四种变换。在进行 RS 分析法时，对载密图像要选取部分像素进行 F_1、F_{-1} 变换，而后比较 R_+ 与 R_-、S_+ 与 S_- 的值。在负载率为 100%，检测选取 50%像素点进行处理时，表 1 给出了对新的 LSB 算法进行 RS 分析的结果。

表 1　对新算法进行 RS 分析时像素值变化情况

隐写时像素情况	RS 分析时 F_1 处理后像素情况	RS 分析时 F_{-1} 处理后像素情况
1/4 像素不变化	1/8 像素不变化	1/8 像素不变化
	1/8 像素进行 F_1 处理	1/8 像素进行 F_{-1} 处理
1/4 像素进行 F_1 处理	1/8 像素进行 F_1 处理	1/8 像素进行 F_1 处理
	1/8 像素不变化	1/8 像素进行 $F_{-1}F_1$ 处理
1/4 像素进行 F_{-1} 处理	1/8 像素进行 F_{-1} 处理	1/8 像素进行 F_{-1} 处理
	1/8 像素进行 F_1F_{-1} 处理	1/8 像素不变化
1/8 像素进行 $F_{-1}F_1$ 处理	1/16 像素进行 $F_{-1}F_1$ 处理	1/16 像素进行 $F_{-1}F_1$ 处理
	1/16 像素进行 $F_1F_{-1}F_1$ 处理	1/16 像素进行 F_1 处理
1/8 像素进行 F_1F_{-1} 处理	1/16 像素进行 F_1F_{-1} 处理	1/16 像素进行 F_1F_{-1} 处理
	1/16 像素进行 F_{-1} 处理	1/16 像素进行 $F_{-1}F_1F_{-1}$ 处理

从上述分析中可以看出，在进行 RS 分析时，F_1 处理和 F_{-1} 处理增加的混乱度大致相似，所以 RS 分析法对该方法将失去作用。

为了验证上述结果，我们选择了一些 256×256 灰度图像进行分析试验，结果如图 3 所示。图中横坐标表示负载率，纵坐标表示 R_+，S_+，R_-，S_- 的值。

从图上可以看出 $R_+ \approx R_-$ 和 $S_+ \approx S_-$，所以 RS 分析法对改进的隐写方案无法进行有效检测。

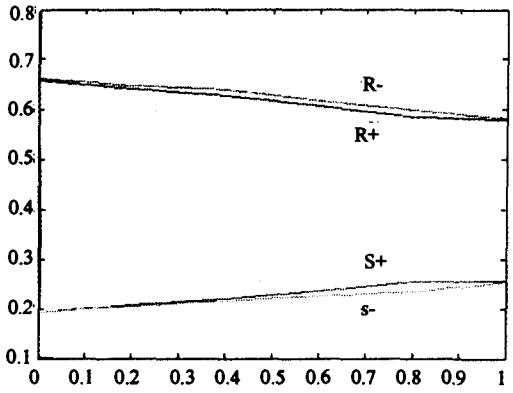

图 3　RS 分析法对新算法的实验结果

(2)抵御统计分析法

统计分析法首先计算截取图像的 R 值，而后对该图像利用 LSB 方法嵌入测试信息，并计算对应的 R' 值，通过比较 R 与 R' 值的大小，判断截取图像是否隐写及隐写信息的量。对于传统的 LSB 隐写，由于灰度值为 $2i$ 的像素中有大约一半改为 $2i+1$，灰度值为 $2i+1$ 的像素中也会有大约一半改为 $2i$，因此在隐写后的载体中，随着嵌入率 α 的增大，h_{2i+1} 与 h_{2i} 趋近相等。而对于改进的算法，期望值如下：

$$E(h_{2i}) = f_{2i} - \frac{3}{4}\alpha f_{2i} + \frac{1}{4}\alpha f_{2i+1} + \frac{1}{4}\alpha f_{2i-1} + \frac{1}{8}\alpha f_{2i+2} + \frac{1}{8}\alpha f_{2i-2}$$

$$E(h_{2i+1}) = f_{2i+1} - \frac{3}{4}\alpha f_{2i+1} + \frac{1}{4}\alpha f_{2i+2} + \frac{1}{4}\alpha f_{2i} + \frac{1}{8}\alpha f_{2i+3} + \frac{1}{8}\alpha f_{2i-1}$$

$$E(h_{2i+2}) = f_{2i+2} - \frac{3}{4}\alpha f_{2i+2} + \frac{1}{4}\alpha f_{2i+3} + \frac{1}{4}\alpha f_{2i+1} + \frac{1}{8}\alpha f_{2i+4} + \frac{1}{8}\alpha f_{2i}$$

所以

$$E(h_{2i+1} - h_{2i}) = (1-\alpha)(f_{2i+1} - f_{2i}) + \frac{1}{8}\alpha(f_{2i+2} - f_{2i-1} + f_{2i+3} - f_{2i-2})$$

$$E(h_{2i+2} - h_{2i+1}) = (1-\alpha)(f_{2i+2} - f_{2i-1}) + \frac{1}{8}\alpha(f_{2i+3} - f_{2i} + f_{2i+4} - f_{2i-1})$$

这样，$E(h_{2i+1} - h_{2i})$ 与 $E(h_{2i+2} - h_{2i+1})$ 都不会随着嵌入率 α 的增大而趋近于 0，F_1、F_2 将不再遵循统计分析法的规律。

进一步的实验分析还得出，当隐秘信息嵌入率小于 40% 时，载体图像的 R 值趋近于 1；当隐秘信息嵌入率小于测试信息嵌入率时，R 总小于 R'；当隐秘信息嵌入率大于等于测试信息嵌入率时，大部分情况下 R 都小于 R'。而统计分析法中，为避免较大的误检率，总是选择较大的测试信息嵌入率，所以当隐秘信息嵌入率小于 40%时，该方法失效。图 4 显示了上述实验结果，图中横坐标隐秘信息嵌入率，纵坐标是 R 与 R' 的值。

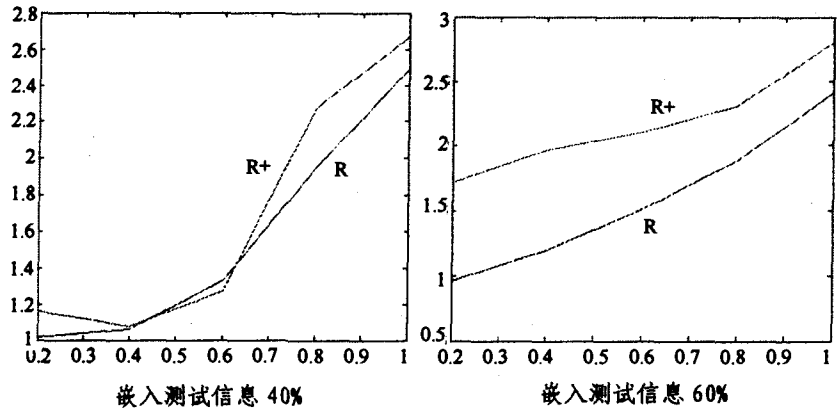

图 4　统计分析法对改进 LSB 算法的实验结果

综上分析，改进的 LSB 隐写算法可以有效地抵御 RS 分析法和统计分析法的攻击，从而提高了传统 LSB 隐写算法的安全性。

结束语　本文在对传统的 LSB 隐写算法进行分析的基

础上，提出了一个改进的 LSB 隐写算法，该算法采用载体图像中最低两位隐藏信息，使用了 F_1、F_{-1}、F_1F_{-1}、$F_{-1}F$ 四种变换。文中还对 RS 分析方法和统计分析法进行了阐述，并利用这两种技术对新算法的安全性进行了分析。理论分析及实验结果表明，改进算法提高了传统 LSB 算法的安全性，可以抵御 RS 分析方法和统计分析法的攻击。

参 考 文 献

1 陈波，谭运猛，吴世忠. 信息隐藏技术综述[J]. 计算机与数字工程，2005，33(2)：21～23
2 Wu D C, Tsai W H. A steganographic method for images by pixel-value differencing[J]. Pattern Recognition Letters, 2003, 24: 1613～1626
3 Zhang X, Wang S. Steganography using multiple-base notational system and human vision sensitivity[J]. IEEE Signal Processing Letters, 2005, 12(1): 67～70
4 Chang C C, Tseng H W. A Steganographic method for digital images using side match[J]. Pattern Recognition Letters, 2004, 25: 1431～1437
5 夏煜，郎荣玲，等. 基于图像的信息隐藏检测算法和实现技术研究综述. 计算机研究与发展，2004，41(14)：728～736
6 Fridrich J, Goljan M, Du R. Detecting LSB steganography in color and gray-scale images[J]. Magazine of IEEE Multimedia (Special Issue on Security), 2001, 8 (4): 22～28
7 Fridrich J, Du R, Long M. Steganalysis of LSB encoding in color Images[R]. ICME'2000, New York, USA, 2000
8 邓倩岚，林家骏. 基于统计的 LSB 隐写分析方法[J]. 计算机安全，2006，1：23～24
9 Westfeld A, Pfitzmann A. Attacks on steganographic systems [J]. Lecture Notes in Computer Science, 1999, 1768: 61～76
10 Ker A D. Steganalysis of LSB matching in grayscale images[J]. IEEE Signal Processing Letters, 2005, 12(6): 441～444

(上接第 89 页)

另外，目前没有任何路由协议声称可以抵御虫洞攻击。如图 5 所示，两个合谋节点 A 之间的虚线表示“虫洞”，合谋节点通过“虫洞”转发来自合法节点的信息，导致 R_1、R_2 误以为互为邻节点。攻击节点运用与中间人攻击同样的手段转发 rrep 报文，同理 EndairA 无法抵御这种攻击，最终 S 会接受 (R_1R_2) 为一条正确路由。但在 EndairALoc 中，源节点 S 检查位置信息列表 $L_D L_{R2} L_{R1}$，L_{R2} 与 L_{R1} 之间距离远远超出节点通信范围，故可知相应 route list 不是正确路由，即 EndairALoc 协议具有抵御虫洞攻击的能力。

5 安全和性能分析

5.1 安全性分析

EndairALoc 除能够抵御中间人攻击以及虫洞攻击，还保持了 EndairA 原有的安全性，分析如下：

(1)恶意结点修改报文中的控制信息及位置信息：控制信息包括身份标识、序列号等，位置信息是节点的位置。这些消息包含在消息认证码的散列范围中，因此任何破坏完整性的操作都将被源节点检查消息认证码时发现。

(2)恶意结点丢弃 rreq 或 rrep 报文：EndairALoc 仍属于改进的 DSR 安全路由协议，一个路由请求将返回多条应答路由，少量的恶意节点将不会影响路由的建立。

(3)重放攻击：恶意结点向网络中传播以前传送过的 rreq 或 rrep 报文，由于报文中携带的序列号保证了消息的新鲜性，这些重放报文将被其它结点作为滞后报文丢弃。

5.2 性能分析

安全路由协议较普通路由协议扩展了安全功能，必然相应地引入一定的网络负载和能耗。源节点消息认证码的验证和位置信息列表的检查增加了源节点的计算开销和路由建立的时延，但因路由请求阶段简单快捷，节点操作步骤少，只有路由应答报文传播过程中节点需计算消息认证码，并且一次路由请求可以获得多条路由信息，因此路由开销整体不大。另外 EndairALoc 采用的是计算量小的对称密钥密码体制的消息认证码，而不是 EndairA 选择的公钥密码体制的数字签名的方式。文[11]对公钥密码算法和对称密码算法的能耗做了量性分析，分析结果如表 1 所示，公钥密码算法比对称密码算法的能耗大几个数量级，因此 EndairALoc 减小了节点计算量，延长了网络寿命。综上可知，EndairALoc 在增强了安全性的同时，并没有带来过多的能耗，更加适合 Ad hoc 网络这种能量受限的网络。

表 1 不同加密算法的能耗

算法	能耗
Public-key(RSA, DSA, ECDSA)	100～500mJ
Secret-key(DES, AES, IDEA)	2～5uJ
Hash(MD5, SHA, HMAC)	0.5～1uJ

小结 本文通过对安全路由协议 EndairA 的分析，指出该协议虽然避免了 Aridane 协议的缺陷，但是并非具有其所声称的能够抵御所有 active-1-y 攻击的能力。我们发现了一种 EndairA 不能抵御的 active-0-1 型攻击——中间人攻击，并据此提出了一种新的安全路由协议 EndairALoc。分析表明该协议不仅保持 EndairA 原有的安全性，而且具有抵御中间人攻击和目前没有任何协议可以抵御的虫洞攻击的能力。另外，由于该协议利用对称密钥机制替代了 EndairA 中采用的公钥签名机制，因此降低了路由建立所需的能耗。未来我们将进一步研究能够抵御更强的敌手模型的新的 Ad hoc 网络安全路由协议。

参 考 文 献

1 Hu Y C, Perrig A. A survey of secure wireless Ad hoc routing. Security & Privacy Magazine, IEEE, 2004, 2: 28～39
2 Papadimitratos P, Haas Z. Secure routing for mobile ad hoe networks. In: Proc. of the SCS Communication Networks and Distributed Systems Modeling and Simulation Conf. San Antonio TX, 2002. 27～ 31
3 Hu Y C, Perrig A, Johnson D B. Ariadne: a secure on-demand routing protocol for Ad hoc networks. In: Proc. of the Eighth ACM Intl Conf. on Mobile Computing and Networking (MOBICOM 2002). Atlanta, GA, 2002. 23～28
4 Zapata M G. Se curing Ad hoc routing protocols. In: Proc. of ACM workshop on wireless Security. Atlanta, Sep. 2002. 1～9
5 Sanzgir K, Dahill B. A secure routing protocol for Ad hoc networks. In: Proc. of the 10 IEEE Intl. Conf. on Network Protocols, 2002. 1～10
6 Ghazizadeh S, Ilghami O, Sirin E. Security—aware adaptive dynamic source routing protocol. In: Proc. of the 27th Annual IEEE Conf. on Local Computer Networks, 2002
7 Hu Y C, Johnson D B, Perrig A. SEAD: secure efficient distance vector routing for mobile wireless Ad hoc networks. Ad hoc Networks, 2003, 1(1): 175～192
8 Johnson D B, Maltz D, Hu Y C. The dynamic source routing protocol for mobile Ad hoc networks. http://www.ietf.org/internet-drafts/draft-ietf-manet-dsr-10.txt, 2005
9 Buttyan L, Vajda I. Towards provable security for Ad hoc routing protocols. In: Proc. of the 2nd ACM Workshop on Security of Ad hoc and Sensor Networks, 2005
10 AdHoc Positioning System (APS). GLOBECOM 2001 - IEEE Global Telecommunications Conference, 2001(1): 2926～2931
11 Potlapally N, Ravi S, Raghunathan A, Jha N. Analyzing the EnergyConsumption of Security Protocols. ISLPED'03, 2003