基于依赖搜索树的电力通信网络告警关联方法的研究

王保义，郭雅薇，史占成，张少敏

（华北电力大学计算机科学与技术学院， 河北 保定 071003）

摘要：从电力通信网的告警机制出发，针对其网络结构和告警信息数据的特征，对告警关联规则及故障定位进行了研究，利用搜索树可以减少搜索空间和覆盖结点的特点提出了一种基于依赖搜索树的告警关联方法。该方法的基本思想是把告警序列中属于一类的告警信息聚合在一起，并用较少的信息代替这一类，从而使海量告警信息简约化，最终能够准确的表达故障信息，达到定位故障的目的。该方法基于故障传输模型，建模简单，能够直接利用结点的依赖关系，并且适用于多故障源的情况。通过算例分析，证明此方法能对电力通信网络的告警信息进行分析，并且能快速准确发现故障源，便于网络的维护。

关键词：告警关联；聚类关联；依赖搜索树；有向图；电力通信网；故障定位；过滤关联

Research of alarm correlation method based on dependency search tree in electric power communication network

WANG Bao-yi, GUO Ya-wei, SHI Zhan-cheng, ZHANG Shao-min

(North China Electric Power University, Baoding 071003, China)

Abstract: Aiming at the framework of electric power communication network and the character of alarm data, this paper investigates the formulae of alarm correlation and fault location. Taking into account dependency search tree which can decrease searching space and nodes of overlay, a method of alarm correlation based on dependency search tree is presented. Alarms belonged to one class in alarm sequence are aggregated, and they are replaced by less information. In this way, large numbers of alarm information are reduced and so the faults can be located accurately. This method based on the model of fault transmission is easy to build a model, can make use of the dependency of nodes directly and it is still applicable when many faults occur. A practical example of faults in electric power communication network is given to prove that this method can analyze alarm information of electric power communication network and locate faults rapidly and exactly. So network can be maintained expediently.

Key words: alarm correlation; clustering correlation; dependency search tree; directed graph; electric power communication network; fault location; filtration correlation

中图分类号： TM76 文献标识码： A 文章编号： 1003-4897(2008)06-0059-06

0 引言

随着电力市场的快速发展，电力调度自动化系统、配电网调度自动化系统、综合自动化系统、继电保护系统、自动抄收表系统、95598 客服系统、办公自动化系统等逐步建立和完善，这些自动化系统都有各种大量的数据、话音和图像业务信息需要通过通信专网来传输。电力通信网是电力系统的专用通信网，是电力市场运行的通信基础，承担的主要任务是传递各种电力生产和管理业务信息。

电力通信网是集传输、交换、终端为一体的有多个环节构成的复杂系统，包括载波、微波、光纤、程控交换、图像监控、电源监控和录音系统等，行业的特殊性对通信网的安全、稳定提出了更高的要求。在电力通信网络中，设备以及线路的故障会以告警信息的形式显示，而对于一个企业网络，每天都会产生数量庞大的告警。如何从海量的告警信息中挖掘出有价值的信息并准确定位电力通信网故障，成为急待解决的问题。文献[1]在电力通信网综合管理系统中引进告警机制，为未来的故障处理提供了理论基础，但相关性分析部分简单。文献[2]提出了一种智能综合告警收集机制，针对某些设备物理层不开放的情况，避开物理层，在协议层实现部分维护端口整合和全部告警系统整合，但没有提及海量告警信息的处理。

本文从电力通信网的告警机制出发，针对其网

络结构和告警信息数据的特征，对告警关联[3, 4]规则及故障定位[5, 6]进行了研究，利用搜索树[7]可以减少搜索空间和覆盖结点的特点提出了一种基于依赖搜索树的告警关联方法，有效地提高电力通信网络故障定位速度，更好的对电力通信网进行监控，满足电力系统安全、稳定、高效生产的需求。

1　电力通信网网络的建模

本文提出的告警关联方法基于故障传输模型，这就要求网络拓扑结构必须已知。假定电力通信网的网络拓扑已由拓扑发现软件获取，并存于数据库中。为了告警关联的需要，将网络拓扑形式化[8]，即将被管网元及端口结点化。

（1）网络拓扑：将网络拓扑抽象为一个有向图 $D=<N, E>$。其中 $N=\{n_1, n_2, n_3, \cdots, n_n\}$，为网络设备结点的集合；$E=\{e_1, e_2, e_3, \cdots, e_n\}$，为链路集合。

（2）设备：设备可以用一个二维向量 $n_i=(n, p_n)$ 表示。其中 n 为设备的标识，随机产生；p_n 为故障概率，数据由厂商提供。

（3）链路：网络中两结点之间的链路表示为一个三维向量 $e_i=(n_i, n_j, C(n_i, n_j))$。其中 n_i 和 n_k 表示链路两端的结点，$C(n_i, n_j)$表示 n_j 依赖于 n_i 的故障概率，即两结点之间的依赖值，数据由历史故障纪录计算得到。

2　电力通信网告警信息的建模

2.1 电力通信网的告警机制

针对电力通信网中设备种类繁多，并且都带有各自的网管，互相独立，互不兼容的特点，本文利用网管协议直接进入设备的网管系统采集告警数据，实现告警信息的采集与整合。网管协议收集到的告警数据的格式见表 1。具体各个字节的含义如表 2 所示。

表 1 告警数据格式

Tab.1 Format of alarm data

B1	B2	B3	B4	B5	B6	B7	B8
7E	81	B3	00	00	00	00	00

2.2 告警数据库设计

为了数据处理的需要，把原始的告警数据转化为统一的标准格式，并存储于电信网告警数据库中[9]。因为每一个告警对应一种告警类型，在这一告警类型中，告警又有一个标准格式，多个标准格式的告警组合在一起构成一个告警序列，因此首先需要把告警类型、告警和告警序列三个参量形式化。

（1）告警类型：告警数据格式中不同的数据对应一种告警类型，用 $G=\{g_1, g_2, g_3, \cdots, g_n\}$表示。

表 2 告警数据位含义

Tab.2 Meanings of alarm data bit

字节顺序	内容说明
第一字节 B1	7E
第二字节 B2	81
第三字节 B3	源地址：0，0，0，ADD4，ADD3，ADD2，ADD1，ADD0
第四字节 B4	REG1：本端设备支路消失告警：LOS16，LOS15，LOS14，LOS13，LOS12，LOS11，LOS10，LOS9
第五字节 B5	REG2：本端设备支路消失告警：LOS8，LOS7，LOS6，LOS5，LOS4，LOS3，LOS2，LOS1
第六字节 B6	REG3：本端线路告警：L—NOP，L—LOF，L—E—3，L—E—6；远端线路告警：R—NOP，R—LOF，R—E—3，R—E—6
第七字节 B7	REC4：远端设备支路消失告警：LOS16，LOS15，LOS14，LOS13，LOS12，LOS11，LOS10，LOS9
第八字节 B8	REG5：远端设备支路消失告警：LOS8，LOS7，LOS6，LOS5，LOS4，LOS3，LOS2，LOS1

（2）告警：给定告警类型 G，告警可以用一个三维向量（A，s，t）表示。其中 $A\in G$，对应于该告警类型；t 为附带的时间戳，是告警发出的时间；s 是发出告警的网络设备标识。

（3）告警序列：在告警类型 G 上的告警序列 S，可以表示为一个三维向量（S，T_s，T_e）。其中 $S=<(A_1, t_1), (A_2, t_2) \cdots (A_n, t_n)>$是由所包含有序告警所组成的序偶，$A_i\in E$ 并且 $t_i\leqslant t_i+1$，$T_s\leqslant t_i<T_e$，$i=1、2、\cdots、n$，T_s 和 T_e 分别为告警序列的起始时间和结束时间。

告警标准化之后，以（A，s，t）的形式存储于数据库中，实例如表 3 所示。

表 3 告警序列实例

Tab.3 Examples of alarm sequence

告警类型 A	告警源 s	告警时间 t
光信号丢失	202.206.212.235	2007-8-1　16：22：15
本端支路 E1 丢失	202.206.212.8	2007-8-1　16：22：39
E—6 误码	202.206.212.221	2007-8-1　16：23：02
对端支路 E1 丢失	202.206.212.39	2007-8-1　16：23：25

3　依赖搜索树模型的构建

3.1 电力通信网告警关联的处理过程

由于网络系统在物理和逻辑上普遍存在着关联性，因此网络中某一点出现故障即可引发一系列相关告警。对于一个大规模电信网来说，大概每天会产生200~10000条告警记录。有时告警也可能只是非故障警告，或者出现许多相似警告，因此这样大量的告警并不能为故障诊断提供精确的信息。

告警关联是管理大量告警信息的基本方法之一。告警关联通过剔除不必要和不相关的信息，减少提交给管理平台或网络操作人员的信息数量，提高了信息的语义表达，能帮助找出产生事件的真正问题或条件，快速定位故障。本文将告警关联的过程分成三步（如图1）：

（1）告警信息预处理：这一过程的目的是去掉非告警的警告数据、合并相同事件、删除相反事件、挑选同源事件中的根源性告警。

（2）聚类关联：指通过提取和抽象网络模型，按照网络中设备结点的依赖关系进行聚类。

（3）过滤关联：指根据聚类关联的结果，按照一定的方法对告警进行一系列的逻辑判断，来完成告警信息的过滤。

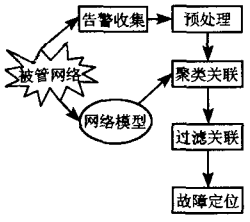

图1 告警关联流程

Fig.1 Process of alarm correlation

3.2 基于依赖关系的覆盖规则

在网络拓扑已知的前提下，电力通信网络中的各种设备的连接关系已知。用二元关系$<n_i, n_j>$表示相互依赖的两个设备结点，则二元组遵守以下的规则：

（1）自反性：对所有的结点 n_i，总存在$<n_i, n_i>$。

（2）对称性：对于同一层设备结点，若$<n_i, n_k>$存在，必有$<n_k, n_i>$存在。对于不同层的结点，则不满足对称性。

（3）传递性：若存在$<n_i, n_k>$和$<n_k, n_j>$，则满足$<n_i, n_j>$。即若 n_k 依赖于 n_i，并且 n_j 依赖于 n_k，则 n_j 间接依赖于 n_i。

基于以上的规则，每一个设备结点都包含直接依赖于它的一组结点，这一组结点叫做这一设备结点的覆盖。覆盖具有如下规则：

（1）网络有向图 $D=<N, E>$中的任意结点 n_i 的覆盖为 n_k；

（2）n_i 可达 n_k，且跳数为1；

（3）n_k 的覆盖中的结点间接依赖于 n_i。

寻找结点的覆盖的目的是将此结点与其覆盖聚合成一类，在结点的搜索过程中，凡是属于此结点覆盖的告警结点均可以用此结点来代替。

3.3 依赖搜索树的构造

随着电力系统通信网与网络技术的融合，电力系统通信网已经由原来多个独立的网络（根据不同的特定业务如行政通信、调度通信、数据通信等，电力系统通信网被设计为多个网络），变为现在的多网合一。因此这样的网络同于非专业网络，设备具有明显的层次关系。根据这样的特点，可以将电力通信网划分成多个子网络，包括高层设备子网（即网络层）和各个特定业务子网（即物理层），结构如图2所示。

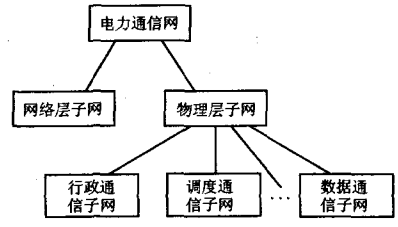

图2 电力通信网子网划分示意图

Fig.2 Sketch map of compartmentalizing subnet in electric power communication network

一个完整的电力通信网被划分成几个子网络之后，可以按照网络的层次关系，先对子网中的告警结点进行关联，在依次向高层设备延伸，最终实现整体的关联。对于各个子网的结点，如何更明显的显示它们之间的关系，是需要解决的问题。按照结点之间的依赖值，将这些结点以一棵生成树的形式表示，即最大依赖值生成树，这样可以更精确的用少量的信息代理大量信息，从而达到告警信息关联的目的。

3.4 在树的先序遍历路径上过滤结点

在结点的搜索过程中，需要对结点排列一个合理的顺序，使告警信息依次经过结点的查找更加有效。由于网络的层次性，以及搜索树中各结点具有的依赖性，树的先序遍历能使结点的层次性更加突出，减少查找次数，提高查找的效率。

对于树上的每一个结点，需要从其覆盖结点中找到待搜索的子结点。为了在每一个结点能够搜索更多的结点向量达到过滤更多结点的目的，并且结点故障最可能引起告警的结点要首先被关联，寻找子结点要遵循下面的规则：

（1）子结点要有大量的覆盖结点；
（2）子结点依赖于父结点的依赖值较大。

4 基于依赖搜索树的告警关联算法

上文中描述的告警关联过程，涉及到了三个算法：依赖搜索树构造算法、聚类关联算法、过滤关联算法，分别在下面的小节中加以介绍。

4.1 依赖搜索树构造算法

给定图 $G=\{N, E\}$，表示电力通信网的子网络。结点 s 为依赖搜索树的根结点，子网中其他结点集合为 D，构造以 s 为根结点且包含所有结点的最大依赖值生成树 T。引进 3 个向量 Md、Mist 和 Parent。Md[u]表示当前搜索到的从源节点 s 到结点 u 的最大依赖值；Parent[u]表示在算法所选择的路径上结点 u 的前一个节点，也就是生成树 T 上结点 u 的父节点；Mist[u]表示生成树 T 上已计算的结点到结点 u 的路径长度。设 ADJ(u)表示结 u 的邻接结点集，Q 为一个待发展节点的序列，存储已被访问但尚未被添加到生成树 T 上的结点。算法如下：

```
DST(G,s,D)
Initialize tree T with source node s and Q;
For all w∈N Do    //算法初始化
If w∈ADJ(s) Then
Dist[w]←C(s,w);
Mist[w]←C(s,w);
Parent [w]←s;
Q←Q∪{w};
Else Dist[w]←∞; Mist[w]←∞; Parent[w]←nil;
End If
End For
Dist[s]←0; Mist[s]←0; Parent[s]←nil;
//源节点为依赖搜索树的根结点
While there exists a node in D that has not been added to tree T Do
Select node u from Q which statisfies Md[u]=max{Md[m]|m∈Q};
If u is a leaf node Then
Establish path from source node s to node u;
Mist[u]←0;          //叶子结点则路径长度为 0
End If
For all w∈ADJ(u) Do
If Md[w]=* Then Q←Q∪{w} End If
If Md[u]*C(u,w) >Md[w] Then
//调整结点的可达最大依赖路径长度
Dist[w]←Md[u]*C(u,w);
Mist[w]←Mist[u]*C(u,w);
Parent[w]←u;
Else If Dist[u]*C(u,w)=Dist[w] Then
//最大依赖路径不唯一
If Mist[u]+1 >Mist[w] Then
//发现更优的父节点
Mist[w]←Mist[u]+1;
Parent[w]←u;      //记录最优父结点
End If
End If
End For
End While
```

4.2 基于依赖搜索树聚类关联算法

假设 N_1 为树 T 的结点，N_T 为 N_1 的覆盖，算法如下：

```
CC(A,T,A1)
p→T;
while(N1)
p→N1;
calculate NT;
  while(NT! =nil)
  Filtrate A' belonged to NT from A;
//从告警序列中找出属于结点覆盖的告警
  A1←A'∪A1;
  Select node N11 from NT;
  calculate NT;
  End While
PT(N,N1)
//按照先序路径寻找树的下一个结点
End While
PT(N,N1)
//按树的先序遍历路径寻找下一个结点
  If N has right child Then
  N1←right child; output N1;
  Else If N has left child Then
  N1←left child; output N1;
  Else If N has brother Then
  N1←brother; output N1;
  Else return
  End If
```

4.3 基于依赖搜索树过滤关联算法

利用聚类关联的函数，将一类告警结点用类中心代替，算法如下：

```
FC(Ai, A')
//输入为类 Ai，输出为过滤后的告警序列 A'
i=1;
for(i=1; i<=n; i++)
```

CC(A,T,A_i);
put A_i into A';　　//用 N_i 代替其中心的一类;
A"←A- A_i;
put A" into A';
//将剩余的告警结点添加到输出中
End For
output A';

5　算例分析

图 3（a）、（b）、（c）所示的是某电力系统的通信网络。某一时间段内，有两个设备发生故障，并且收集到这一时间段内的告警信息，现按照算法确定故障源。

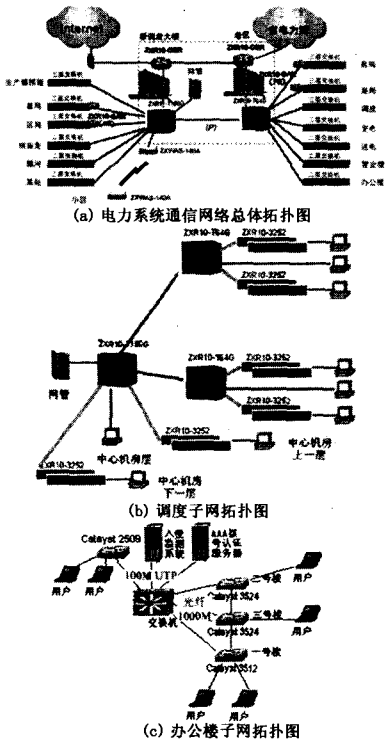

图 3　电力通信网络拓扑图

Fig.3 Typology of power communication network typology

网络的拓扑是已知的，即设备间的连接关系已知。首先，进行网络建模，根据设备的依赖关系将网络拓扑抽象为依赖图，设备抽象为结点，如图 4 所示（其中无向边表示两结点相互依赖）。查询网络拓扑的数据库，得到各结点之间的依赖值，下面给出一个子网的结点依赖值（其它结点略），分别为 C(1,2)=C(2,1)=0.050，C(1,3)=0.045，C(2,4)=0.037，C(3,4)= C(4,3)=0.025，C(3,5)=0.048，C(3,6)=0.029，C(3,7)=0.026，C(3,8)=0.02，C(3,9)=0.035，C(3,10)=0.019，C(4,11)=0.031，C(4,12)=0.033，C(4,13)=0.016，C(4,14)= 0.005，C(4,15)=0.002，C(4,16)=0.011，C(4,17)=0.009。然后按照设备的属性将整个网络划分成子网，根据算法 5.1，将每个子网构造成依赖生成树，如图 5 所示。将告警信息整理并标准化，分别是结点 2、3、4、7、9、11、17、19、22、24、30、37、41 发出告警。

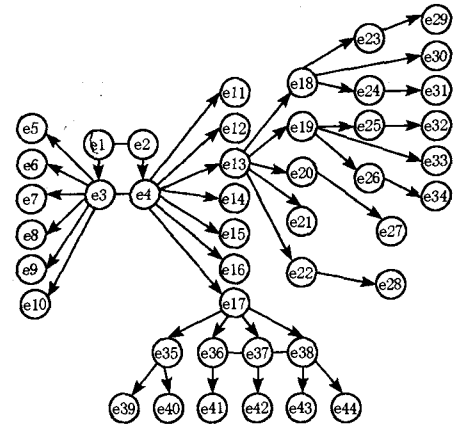

图 4　电力通信网络依赖图

Fig.4 Electric power communication network dependency graph

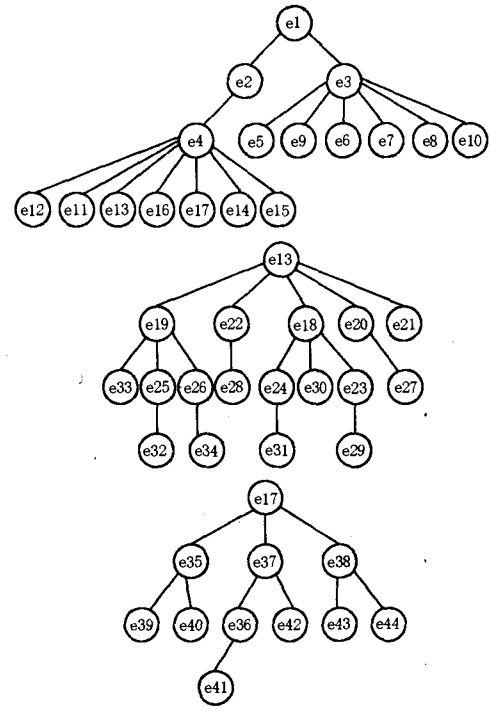

图 5　层次子树

Fig.5 Hierarchical subtree

接下来，按照算法 4.2 对每棵树上的结点进行

筛选。由于自底向上的顺序可以减少搜索空间，提高算法的效率，因此从低层的子网开始搜索。这样。从第二棵树开始，初始时一类中只有中心结点 13，结点 13 的覆盖为结点 18、19、20、21、22，其中 19、22 在告警序列中，将它们加入到此类中。然后从覆盖中挑选依赖值大的结点 19 作为下一个分量结点继续筛选，此结点搜索完之后，按照先序遍历的路径搜索下一个结点。依此类推，直到结点覆盖为空，此棵树搜索完毕，得到以结点 13 为中心的一类为{13、19、22、24、30、37}。调用算法 4.3 将告警序列中的{19、22、24、30、37}用 13 代替，并开始下一棵树的搜索。第三棵树搜索完毕后得到{17、41}，调用算法 4.3 将告警序列中的结点 41 过滤掉。最后完成第一棵树的搜索，得到两类{1、2、3、4、11、17}、{3、7、9}，这样告警序列就可以简化为结点 1、13，而它们就是真正的故障源。

6 结论

本文提出了一种基于依赖搜索树的告警关联新方法，针对电力通信网的网络结构和告警信息的特征，把告警序列中属于一类的告警信息聚合在一起，并用较少的信息代替这一类，从而使海量告警信息简约化，最终能够准确地表达故障信息，达到定位故障的目的。算例分析证明该方法能有效地分析告警信息，确定故障源，从而便于电力通信网的维护。

参考文献

[1] 李立达.电力通信网综合管理系统故障告警机制的设计与实现[J].电力系统通信，2006，27(17)：13-16.
LI Li-da. Design and Realization of Alarm Mechanism of Fault in Power Communication Network Integrated Management System[J]. Telecommunications for Electric Power System, 2006,27(17):13-16.

[2] 李良城，苏建华.智能综合网络告警收集系统在电力通信系统中的应用[J].四川电力技术，2006，29(4)：81-84.

[3] 孙朝晖，张德运，李庆海.网络故障管理中的自动告警关联[J].计算机工程，2004，30(5)：30-34.
SUN Zhao-hui，ZHANG De-yun，LI Qing-hai. Automated Alarm Correlation in Network Fault Management[J].Computer Engineering，2004,30(5):30-34.

[4] 杨洪涛，王继龙.网络事件管理系统中关联技术的选择及实现[J].计算机工程，2006，32(4)：197-199.
YANG Hong-tao, WANG Ji-long. Selection of Correlations Schemes in Network Event Management System[J]. Computer Engineering，2006,32(4): 197-199.

[5] Katzela I, Schwartz M. Schemes for Fault Identification in Communication Networks[A].In:IEEE/ACM Transactions on Network[C].1995.

[6] Choi J，Choi M，Lee S H. An Alarm Correlation and Fault Identification Scheme Based on OSI Managed Object Classes[J].IEEE, 1999.

[7] 肖政，王建新，侯紫峰，等.基于搜索树的告警高效聚类算法和 Bayes 分类器的设计和研究[J].计算机科学，2006，33(8)：190-194.
XIAO Zheng，WANG Jian-xin，HOU Zi-feng，et al. Design and Research of an Alert Clustering Algorithm Based on Search Tree and an Alert Classified Method Based on Bayesian Classifier[J]. Computer Science，2006,33(8):190-194.

[8] 刘缵武.应用图论[M].北京：国防科技大学出版社，2006.
LIU Zan-wu. Application of Graph Theory[M]. Beijing: National University of Defence Technology，2006.

[9] 单莘，朱永宣，郭军.电信网告警数据库中的增量式挖掘技术研究[J].计算机应用研究，2006，(3)：257-260.
SHAN Xin，ZHU Yong-xuan，GUO Jun. Study on Technique of Incremental Mining in Telecommunication Network Alarm Databases[J]. Application Research of Computers，2006,(3): 257-260.

[10] Hains J，Ryder D K，Tinnel L,et al. Validation of Sensor Alert Correlators[J]. IEEE Security & Privacy，2003: 46-56.

[11] Julisch K. Mining Alarm Clusters to Improve Alarm Handling Efficiency[A]. In:17th Annual Computer Security Applications Conference(ACSAC'O1)[C].New York: 2001.

[12] Julisch K，Dacier M. Mining Intrusion Detection Alarms for Actionable Knowledge[A].In:Proc.8th ACM Int1.Conf.on Knowledge Discovery and Data Mining[C]. Edmonton:2002

收稿日期：2007-08-28；　　**修回日期**：2007-10-13

作者简介：

王保义(1964-)，男，教授，研究方向为电力系统自动化、网络与信息安全、信息系统；E-mail: wangbaoyi@126.com

郭雅薇(1982-)，女，硕士研究生，研究方向为数据库与管理信息系统；

史占成(1981-)，男，硕士研究生，主要研究方向为电子商务与政务、网络安全。