基于覆盖广义粗集的模糊性*)

魏 荣[1] 刘保仓[2] 史开泉[1]
(山东大学数学与系统科学学院 济南 250100)[1] (黄淮学院数学科学系 河南 驻马店 463000)[2]

摘 要 本文定义了基于覆盖广义粗集的模糊性度量,给出一种度量表示并讨论此种模糊性度量的性质;定义了正负域覆盖广义粗集的模糊度,并通过一个具体实例给出直观解释。

关键词 广义粗集,覆盖,覆盖约简,模糊性

Fuzziness Based on Covering Generalized Rough Sets

WEI Rong[1] LIU Bao-Cang[2] SHI Kai-Quan[1]
(School of Mathematics and System Sciences, Shandong University, Jinan 250100)[1]
(Department of Mathematics, Huanghuai University, Zhumadian 463000)[2]

Abstract A measure of fuzziness based on covering generalized rough sets is proposed and some characters of this measure are discussed. A measure of fuzziness based on positive-negative region covering generalized rough sets is also put forward. Using a specific example, visual analysis of the new definition is given.

Keywords Generalized rough sets, Covering, Covering reduct, Fuzziness

1 引言

粗集理论[1]是Z Pawlak提出的处理含糊不确定知识的数学理论,它能有效地分析和处理不精确、不一致、不完整等各种不完备信息并从中发现隐含的知识,揭示潜在的规律。模糊理论[2]是L. A. Zadeh提出的刻画模糊现象或模糊概念的数学理论。所谓模糊现象和模糊概念,就是指没有严格的界限划分而使得很难用精确的尺度来刻画的现象和概念。

模糊理论和粗集理论在处理不确定性和不精确性问题方面都推广了经典集合论,它们都可以用来描述知识的不精确性和不完全性,但它们的出发点和侧重点不同。模糊理论考虑的是集合边界的病态定义,与元素和集合的关系相联系,是经典集合论中属于和不属于的推广;粗糙理论考虑的是元素(或对象)间的不可区分性,与集合上的等价关系相联系,是经典集合由等价关系而得出的近似。虽然模糊理论和粗集理论特点不同,但它们之间有着很密切的联系,有很强的互补性。两者的着眼点和计算方法是不一样的,因此不能相互替代,而是必须结合起来使用。许多学者将两种理论结合得到粗模糊集和模糊粗集,也定义了模糊集的粗糙性[3]以及粗集的模糊性度量[4],这些结果广泛应用在模式识别以及智能分析等诸多领域中。

文[4,5]讨论的是Z Pawlak等价关系下粗集的模糊性。许多学者对粗集进行推广,其中覆盖广义粗集[6]被认为是在数据挖掘中具有广泛应用前景的一种模型。本文讨论了基于覆盖广义粗集的模糊性度量问题,给出一种度量表示并分析了此种模糊性度量的性质;定义了基于正负域覆盖广义粗集的模糊度,并通过一个具体实例给出直观解释。

2 覆盖广义粗集的基本概念[6,7]

定义 2.1 设U为论域,C是U的非空子集族,且$\bigcup C=U$,则称集族C为U上的一个覆盖。

显然,U的一个划分也是一个覆盖,覆盖是划分的推广。

下面均假设U为有限论域,C为U上的有限覆盖,即由有限个子集构成的覆盖。

定义 2.2 设U为有限论域,C为U上的一个覆盖,$\langle U,C\rangle$称为覆盖近似空间。

定义 2.3 设$\langle U,C\rangle$为一覆盖近似空间,$x\in U$,称下式为x的最小描述:

$MD(x)=\{K\in C|x\in K\wedge(\forall x\in S\wedge S\subseteq K\Rightarrow K=S)\}$。

定义 2.4 对任意$X\subseteq U$,集族$C_*(X)=\{(K\in C|K\subseteq X\}$称为$X$的覆盖下近似集族。

集合$X_*=\bigcup C_*(X)$称为X的覆盖下近似集。

集合$X_*^*=X-X_*$称为X的覆盖边界集。

集族$Bn(X)=\{Md(x)|x\in X_*^*\}$称为$X$的覆盖边界近似集族。

集族$C^*(X)=C_*(X)\bigcup Bn(X)$称为$X$的覆盖上近似集族。

集合$X^*=\bigcup C^*(X)$称为X的覆盖上近似集。

当$C^*(X)=C_*(X)$时,称X相对于覆盖C是可定义的,否则称X相对于覆盖C是不可定义的。

定义 2.5 设C是论域U的一个覆盖,$K\in C$,若K可表示为$C-\{K\}$中集合之并,则称K为C的可约元,否则为不可约元。

定义 2.6 设C是论域U的覆盖,若C中每个元均不可约,则称C是不可约的;去掉C中所有可约元,得到的仍是U的一个覆盖,称为C的约简,记为$reduct(C)$。

定义 2.7 设A为论域X中的模糊集合,$\lambda\in[0,1]$,定义A的λ-截集为

$A_\lambda=\{x\in X:\mu_A(x)\geqslant\lambda\}$

*)山东省自然科学基金资助项目(Y2004A04)。**魏 荣** 硕士生,主要研究领域为粗集、模糊集理论及其应用;**刘保仓** 副教授,主要研究领域为粗集理论及其应用;**史开泉** 教授,博导,主要研究领域为粗集、模糊集理论及其应用。

3 基于覆盖广义粗集的模糊度

定义 3.1[8] 设$U=\{x_1,x_2,\cdots x_n\}$，A为论域U中的模糊集合，$\forall x\in U$，令$A_{0.5}(x_i)=\begin{cases}1,\mu_A(x_i)\geqslant 0.5\\0,\mu_A(x_i)<0.5\end{cases}$，记$d_P(A)=\frac{2}{n^{1/P}}(\sum_1^n|\mu_A(x_i)-A_{0.5}(x_i)|^P)^{1/P}$，其中$P>0$，则称$d_P(A)$为模糊集$A$的模糊度，特别地$P=1,2$分别称为 Hamming 模糊度$d_1(A)$和 Euclidean 模糊度$d_2(A)$。

定义 3.2 设$\langle U,C\rangle$为一覆盖近似空间，C的约简记为$reduct(C)=\{K_1,K_2,\cdots,K_m\}$，$X\subseteq U$，则由$X$和$C$确定了一个$U$上的模糊集：

$F_X^C=\{(x,\mu_{F_X^C}):x\in U,$

$\mu_{F_X^C}(x)=\max\{\frac{|K_i\cap X|}{|K_i|},x\in K_i\}\}$

例 1 $U=\{a,b,c,d\}$，$C=\{K_1,K_2,K_3,K_4,K_5\}$，其中：

$K_1=\{a,b\}$，$K_2=\{a,c\}$，$K_3=\{c,d\}$，$K_4=\{d\}$，$K_5=\{a,c,d\}$。若$X=\{b,c\}$，则

$reduct(C)=\{K_1,K_2,K_3,K_4\}$；

$$\mu_{F_X^C}(a)=\max\{\frac{|K_1\cap X|}{|K_1|},\frac{|K_2\cap X|}{|K_2|}\}=\max\{\frac{1}{2},\frac{1}{2}\}=\frac{1}{2};$$

$$\mu_{F_X^C}(b)=\max\{\frac{|K_1\cap X|}{|K_1|},\frac{|K_3\cap X|}{|K_3|}\}=\max\{\frac{1}{2},\frac{1}{2}\}=\frac{1}{2};$$

$$\mu_{F_X^C}(c)=\max\{\frac{|K_2\cap X|}{|K_2|},\frac{|K_3\cap X|}{|K_3|}\}=\max\{\frac{1}{2},\frac{1}{2}\}=\frac{1}{2};$$

$$\mu_{F_X^C}(d)=\max\{\frac{|K_3\cap X|}{|K_3|},\frac{|K_4\cap X|}{|K_4|}\}=\max\{\frac{1}{2},0\}=\frac{1}{2};$$

$$d_1(F_X^C)=\frac{2}{4}(\sum_{x_i\in U}|\mu_{F_{X^c}}(x_i)-(F_X^C)_{0.5}(x_i))|=1;$$

$$d_2(F_X^C)=\frac{2}{\sqrt{4}}(\sum_{x_i\in U}|\mu_{F_{X^c}}(x_i)-(F_X^C)_{0.5}(x_i))|^2)^{1/2}=1。$$

下面讨论基于覆盖广义粗集的模糊度的性质。

性质 3.1 设$\langle U,C\rangle$为一覆盖近似空间，下述各式成立：

(1)$d_1(F_U^C)=0$； (2)$d_2(F_U^C)=0$；

(3)$d_1(F_\phi^C)=0$； (4)$d_2(F_\phi^C)=0$。

证明：仅证(1)式，其余证明类似可得。

$$\forall x_i\in U,\mu_{F_U^C}(x_i)=\max\{\frac{|K_i\cap U|}{|K_i|},x_i\in K_i\}=\max\{\frac{|K_i|}{|K_i|},x_i\in K_i\}=1。$$

故$(F_X^C)_{0.5}(x_i)=1$，从而易得$(f_U^C)_1=0$。

性质 3.2 覆盖近似空间的可定义集的模糊度为 0。

证明：若X为覆盖近似空间$\langle U,C\rangle$上的可定义集，则$Bn(X)=\{Md(x)|x\in X_*^*\}=\phi$。$\forall x\in X$，必有$\mu_{F_X}(x)=1$，否则，$x\in X_*^*$，从而$Bn(X)$不空，矛盾。从而$d_P(F_X^C)=0$。

性质 3.3 设$\langle U,C\rangle$为一覆盖近似空间，$X\subseteq U$，$\forall x\in U$，则有：

$\mu_{F_X^C}(x)+\mu_{F_{\sim X}^C}(x)\geqslant 1$。

证明：$\forall x\in U$，

$$\mu_{F_X^C}(x)+\mu_{F_{\sim X}^C}(x)=\max\{\frac{|K_i\cap X|}{|K_i|},x_i\in K_i\}+\max\{\frac{|K_i\cap\sim X|}{|K_i|},x_i\in K_i\}\geqslant\frac{|K_i\cap X|}{|K_i|}+\frac{|K_i\cap\sim X|}{|K_i|}=1,x_i\in K_i。$$

这一性质与 Z Pawlak 粗集的模糊度性质不同，即覆盖广义粗集的模糊度不满足$(F_X^C)^c=F_{\sim X}^C$。

性质 3.4 设$\langle U,C\rangle$为一覆盖近似空间，$X\subseteq U$，$Y\subseteq U$，则下述各式成立：

(1)$F_{X\cup Y}^C\supseteq F_X^C\cup F_Y^C$。

(2)若$X\subseteq Y$或$Y\subseteq X$，则有$F_{X\cup Y}^C=F_X^C\cup F_Y^C$。

证明：(1) $\forall x\in U$，

$$\mu_{F_{X\cup Y}^C}=\max\{\frac{|K_i\cap(X\cup Y)|}{|K_i|},x_i\in K_i\}\geqslant\max\{\frac{|K_i\cap X|}{|K_i|},x_i\in K_i\}=\mu_{F_X^C}。$$

同理$\mu_{F_{X\cup Y}^C}\geqslant\mu_{F_Y^C}$，从而(1)得证。

(2)易证。

性质 3.5 设$\langle U,C\rangle$为一覆盖近似空间，$X\subseteq U$，$Y\subseteq U$，则下述各式成立：

(1)$F_{X\cap Y}^C\subseteq F_X^C\cap F_Y^C$。

(2)若$X\subseteq Y$或$Y\subseteq X$，则有$F_{X\cap Y}^C=F_X^C\cap F_Y^C$。

性质 3.5 的证明与性质 3.4 类似，略去。

4 基于正负域覆盖广义粗集的模糊度

定义 4.1[9] 设$\langle U,C\rangle$为一覆盖近似空间，$\forall X\subseteq U$，分别称

$R^+(X)=\bigcup\{K\in C|K\subseteq X\}$；

$R^-(X)=\bigcup\{K\in C|K\cap X\neq\varnothing\}$，

为集合X关于覆盖C的覆盖正域和覆盖负域，称$(R^+(X),R^-(X))$为X关于覆盖C的正负域覆盖广义粗集。

如果存在$K_i\in C$，$i=1,2,\cdots,n$，使得$X=\sum_{i=1}^n K_i$，则称X为关于覆盖C可正定义的。如果存在$K_i\in C$，$i=1,2,\cdots,m$，使得$\sim X=\sum_{i=1}^m K_i$，则称X为关于覆盖C可负定义的。如果X既可正定义又可负定义，则称X关于覆盖C可全定义。

对应地我们可以定义如下的模糊集：

定义 4.2 设$\langle U,C\rangle$为一覆盖近似空间，C的约简记为$reduct(C)=\{K_1,K_2,\cdots,K_m\}$，$X\subseteq U$，则由$X$和$C$确定了一个$U$上的模糊集：

$$F'^C_X=\{(x,\mu_{F'^C_X}):x\in U,\mu_{F'^C_X}(x)=\begin{cases}\max\{\frac{|K_i\cap X|}{|K_i|},x\in K_i\},(\forall\frac{|K_i\cap X|}{|K_i|}\neq 0)\\0,(\exists\frac{|K_i\cap X|}{|K_i|}=0,x\in K_i)\end{cases}$$

对于上例由定义 4.2 计算：

$$\mu_{F'^C_X}(a)=\max\{\frac{|K_1\cap X|}{|K_1|},\frac{|K_2\cap X|}{|K_2|}\}=\max\{\frac{1}{2},\frac{1}{2}\}=\frac{1}{2};$$

$$\mu_{F'^C_X}(b)=\max\{\frac{|K_1\cap X|}{|K_1|},\frac{|K_3\cap X|}{|K_3|}\}=\max\{\frac{1}{2},\frac{1}{2}\}=\frac{1}{2};$$

$$\mu_{F'^C_X}(c)=\max\{\frac{|K_2\cap X|}{|K_2|},\frac{|K_3\cap X|}{|K_3|}\}=\max\{\frac{1}{2},\frac{1}{2}\}=\frac{1}{2};$$

$\mu_{F'^C_X}(d)=0$；

$$d_1(F'^C_X)=\frac{2}{4}(\sum_{x_i\in U}|\mu_{F'^C_X}(x_i)-(F'^C_X)_{0.5})|=0.75;$$

$$d_2(F'^{C}_{X})=\frac{2}{\sqrt{4}}(\sum_{x_i\in U}|\mu_{F'^{C}_{X}}(x_i)-(F^{C}_{X})_{0.5})|^2)^{1/2}=$$
0.2743。

类似前面讨论的基于广义覆盖粗集的模糊度的性质，基于正负域广义覆盖粗集的模糊度也具有类似性质，在此不再赘述。

结束语 本文讨论了基于覆盖广义粗集的模糊性度量问题，给出一种度量表示并给出此种模糊性度量的性质；定义了正负域覆盖广义粗集的模糊度，并通过一个具体实例给出了直观解释。

参考文献

1 Pawlak Z. Rough Sets. International Journal of Computer and Information Sciences, 1982, 11: 341~356

2 Zadeh L A. Fuzzy Sets. Information and control, 1965, 8(3): 338~356

3 Banerjee M, Pal S K. Roughness of a fuzzy set. Information Sciences, 1996, 93: 235~246

4 Chakrabarty K, Biswas R, Nanda S. Fuzziness in rough sets. Fuzzy sets and systems, 2000, 110: 247~251

5 李兵，吴孟达. 粗糙集研究中的模糊方法. 模糊系统与数学，2002, 16(2): 69~73

6 Bigniew Z, Bonikowski, Bryniarski E. Extension and intension in the rough set theory. Information Sciences, 1998,107: 149~167

7 Zhu W, Wang F Y. Reduction and axiomization of covering generalized rough sets. Information Sciences, 2003, 152: 217~230

8 刘寅普，吴孟达. 模糊理论及应用. 长沙：国防科技大学出版社，1998

9 薛佩军，管延勇. 正负域覆盖广义粗集及其运算公理化. 计算机工程与应用，2005, 41(27):35~37

（上接第 51 页）

个都成立，则新节点相信自身的 IP 临时地址重复，使 IP 地址自加，重复以上地址申请过程。地址分配的流程图如图 2 所示。

当两个网络发生合并的时候，也会出现地址冲突的现象，可以按照 Weak DAD[3] 方法发现地址冲突，当地址重复的两个节点收到 Address_Conflict 消息后，为了防止两个节点同时改变自己的地址，每个节点会用自己发出的地址冲突消息里的时间戳与接收到的地址冲突消息里的时间戳相对比，如果发现收到的时间戳小于自己的时间戳，说明对方先发现自己，那么就要重新选择地址。

3 安全性及效率分析

3.1 安全性分析

本方案能有效防止 ad hoc 网络里的各种相关安全隐患，分别从以下几个方面进行分析。

3.1.1 地址欺骗攻击和地址冲突攻击

在身份验证阶段，由于恶意节点无法得到节点的私钥和合法的数字证书，每个收到请求身份认证的节点都会用证书里的公钥来检查数字签名，用来验证发送者的真实身份。恶意节点就没有机会伪装其它节点的地址或发出假的地址冲突消息。在地址分配阶段，也会用节点数字证书来验证数字签名，这样可以在节点通过了身份认证后，地址分配前的时间段内有效地防止被恶意节点使用地址欺骗攻击和地址冲突攻击。

3.1.2 防止地址耗尽攻击

地址分配过程中，因为消息是加密的，恶意节点无法看到消息里面封装序列号和时间戳，无法猜测出下一个包的序列号是多少，也可以通过判断定时器是否超时，有效地防止了地址耗尽攻击。

3.1.3 防止 Wiener 攻击

Wiener 也叫加拿大人攻击，在身份认证阶段，A 重复使用他用来提问的一次性随机数 R_A 来完成认证，这样可以防止 Wiener 攻击，使得入侵者不能作为中间人来发起攻击[7]。

3.2 效率分析

3.2.1 身份认证资源问题

建立一次身份认证和地址分配的开销比较大。密钥的注册和证书颁发，每个相邻节点对消息进行认证和 RSA 计算的过程，以及计算散列函数，都需要大量的开销。要做到 ad hoc 网络的安全，这种开销几乎是不可避免的。

3.2.2 地址冲突的可能性

在地址分配阶段，采用了对请求者的公钥进行哈希散列，生成 IP 地址的方法。在 IPv4 的空间里，当节点个数达到 500 个时，经生日判定（birthday paradox）计算：地址空间为 C，n 为节点个数。

$$P=1-\frac{C!}{C^n(C-n)!}$$

冲突的几率只有 1.0E-05。而在 IPv6 的空间里，几率则会更小，这样小的冲突几率完全可以满足实际的需要。

结论 本文首先提出了一种多层安全模型，然后提出一种新的安全地址自动分配场景，能有效防止各类相关攻击，为移动自组网实用化打下坚实基础。在本方案中，首先提出一个安全的身份认证场景，新加入节点发出认证请求，经过它与相邻节点的相互认证，最终使新节点通过身份认证。其次就是提出一个安全的地址分配场景，即当新节点通过身份认证后，用散列函数把自己的私钥生成请求地址，经过地址冲突和安全检测后，最终加入网络，可以有效地防止恶意节点对消息的截取、修改、伪造。

参考文献

1 Mesargi S, Prakash R. MANETconf: Configuration of Hosts in a Mobile Ad Hoc Network[A]. In: Proceedings of the IEEE Conference on Computer Communications (INFOCOM) [C]. New York, NY, 2002

2 Zhou H, Ni LM, Mutka M W. Prophet Address Allocation for Large Scale MANET[A]. In: Proceedings of IEEE INFOCOM 2003[C]. San Francisco, CA, 2003

3 Vaidya N H. Weak Duplicate Address Detection in Mobile Ad Hoc Networks". MobiHoc2002, June 2002

4 Hu Y, Perrig A, Johnson D B. Ariadne: a secure on demand routing protocol for ad hoc networks. In: Proceedings of the Eighth Annual International Conference on Mobile Computing and Networking (Mobicom 2002), ACM, Atlanta, GA, September 2002. 12~23

5 Hu Y, Perrig A, Johnson D B. SEAD: secure efficient distance vector routing for mobile ad hoc networks. In: Proceedings of the 4th IEEE Workshop on Mobile Computing Systems and Applications(WMCSA 2002), IEEE, Calicoon, NY, June 2002

6 Sanzgiri K, Dahill B, Levine B N, et al. A secure routing protocol for ad hoc networks. In: International Conference on Network Protocols(ICNP), Paris, France, November 2002

7 Diffie W, Van Oorschot P C, Wiener M. Authentication and authenticated key exchanges. Designs, Codes and Cryptography, 1992, 2:107~125