

控制与决策

Control and Decision

基于信息融合的罪案关联证据推理模型

徐祖润, 刘思峰, 方志耕, 吴洁, 刘勇

引用本文:

徐祖润, 刘思峰, 方志耕, 等. 基于信息融合的罪案关联证据推理模型[J]. *控制与决策*, 2020, 35(1): 228–234.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2018.0455>

您可能感兴趣的其他文章

Articles you may be interested in

平板导电结构缺陷脉冲涡流和超声复合检测方法

Pulsed eddy current and ultrasonic complex testing method for defect detection of planar conductive structures

控制与决策. 2019, 34(4): 743–750 <https://doi.org/10.13195/j.kzyjc.2017.1267>

一种基于证据推理的装备保障资源评估方法

An evaluation method of equipment support resources based on evidential reasoning

控制与决策. 2018, 33(6): 1048–1054 <https://doi.org/10.13195/j.kzyjc.2017.0288>

基于证据马尔可夫随机场模型的图像分割

Image segmentation based on evidential Markov random field model

控制与决策. 2017, 32(9): 1607–1613 <https://doi.org/10.13195/j.kzyjc.2016.0901>

区间信息类型下的突发事件动态预警方法

Dynamic early-warning method of emergency event with interval information

控制与决策. 2017, 32(7): 1306–1312 <https://doi.org/10.13195/j.kzyjc.2016.0731>

考虑可靠性与重要性的证据补偿协调融合方法

Compensation coordinated rule for fusing evidences by considering reliability and importance

控制与决策. 2016, 31(9): 1623–1630 <https://doi.org/10.13195/j.kzyjc.2015.1039>

基于证据推理的联合故障检测方法

Method of joint fault detection based on evidential reasoning

控制与决策. 2016, 31(9): 1589–1593 <https://doi.org/10.13195/j.kzyjc.2015.0989>

基于证据理论-DEA交叉效率的混合型多属性决策方法

Method for hybrid multiple attribute decision-making based on Dempster-Shafer theory and cross efficiency of DEA

控制与决策. 2016, 31(5): 943–948 <https://doi.org/10.13195/j.kzyjc.2015.0277>

考虑应急方案总体优势度的决策方法

Decision method for emergency alternative with considering total superiority degree

控制与决策. 2015(7): 1239–1244 <https://doi.org/10.13195/j.kzyjc.2014.0661>

基于信息融合的罪案关联证据推理模型

徐祖润^{1,2†}, 刘思峰¹, 方志耕¹, 吴 洁³, 刘 勇⁴

(1. 南京航空航天大学 经济与管理学院, 南京 210016; 2. 江苏科技大学 理学院, 江苏 镇江 212003;
3. 江苏科技大学 经济管理学院, 江苏 镇江 212003; 4. 江南大学 商学院, 江苏 无锡 214122)

摘 要: 针对传统犯罪案件中出现的冲突证据难以处理的情况, 提出一种基于证据可信度和灰色关联的冲突证据融合处理方法; 基于相似性视角的灰色关联度作为证据间的联系, 采用改变证据源的证据组合规则来衡量证据源中各个证据之间的贴进度; 考虑到证据的可信度, 提出新的基于证据贴进度的权重确定方法, 以证据的可信度作为权重, 对参与融合的证据的基本概率分配函数进行加权平均, 使证据融合收敛速度更快, 并提升融合效果. 最后以安徽省某市的入室盗窃犯罪案件为例, 运用基于信息融合的关联证据推理方法处理案件中的冲突证据问题, 验证了所提出方法的合理性和有效性.

关键词: 证据理论; 信息融合; 灰色关联; 罪案推理; 冲突证据; 可信度

中图分类号: N941.5

文献标志码: A

Crime-related evidence reasoning model based on information fusion

XU Zu-run^{1,2†}, LIU Si-feng¹, FANG Zhi-geng¹, WU Jie³, LIU Yong⁴

(1. College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China; 2. Faculty of Science, Jiangsu University of Science and Technology, Zhenjiang 212003, China; 3. College of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang 212003, China; 4. School of Business, Jiangnan University, Wuxi 214122, China)

Abstract: For the situations that it's difficult to deal with the conflicting evidences in traditional criminal cases, a conflicting evidences fusion processing method based on evidence credibility and grey incidence is proposed. Based on the degree of grey incidence on similarity perspective as a link between evidences, a combination of evidences that changes the source of evidences is used to measure the closeness between the various evidence sources. Considering the credibility of evidences, a new weight determination method is proposed based on the closeness of evidences. Taking the credibility of the evidence as the weight, the weighted average of the basic probability assignment function of the evidence participating in the fusion is used to make the evidence fusion convergence faster and improve the fusion effect. Finally, through the case of a theft crime in a city of Anhui province, the evidence reasoning method based on information fusion is used to deal with the conflicting evidences in the case, which reflects the rationality and effectiveness of the proposed method.

Keywords: evidence theory; information fusion; grey incidence; crime reasoning; conflicting evidences; credibility

0 引 言

传统罪案证据推理模式属于日常思维模式, 存在着自发性、粗放性、跳跃性等弊端, 这使得冤错案件时有发生. 罪案发生是一种极为敏感和复杂的社会现象, 在社会生活中具有较大的影响, 而冤假错案会将司法机关推到舆论的风口, 使司法的严肃性遭遇到

严峻的挑战. 在当今的信息化时代, 相关机构力图通过有效的量化数据分析方法(包括罪案定量推理)来预防控制和打击犯罪, 但由于犯罪现象自身的不确定性和复杂性, 导致对犯罪问题的量化推理分析和研究非常困难. 由此可见, 罪案推理研究是关系到司法权威性和公信力, 关系到社会公正的世界性难题, 加强

收稿日期: 2018-04-13; 修回日期: 2018-09-28.

基金项目: 国家自然科学基金项目(71671091, 71503103, 11602098); 欧盟第 7 研究框架玛丽·居里国际人才引进计划 Fellow 项目(FP7-PIIF-GA-2013-629051); 中央高校基本科研业务费专项基金项目(NP2015208); 国家级教学团队基金项目(10td128); 国家社会科学基金项目(17BGL028, 17BFX022); 江苏省自然科学基金项目(BK20150157); 江苏省社科基金项目(13GLD019, 14GLC008); 江苏科技大学国家自然科学基金预研项目(2014SL111J); 江苏省普通高校研究生科研创新计划项目(KYZZ_0094).

责任编辑: 李国齐.

†通讯作者. E-mail: xuzurun@163.com.

对罪案证据推理方法在技术层面上的探索是十分必要的^[1]。

目前,针对犯罪证据的研究日趋活跃,Kennedy等^[2]以风险评估和空间分析技术为手段,将分析的实际结果与犯罪空间问题联系起来,探讨了证据与环境等因素对于区分不同犯罪率区域的影响;Johnson等^[3]以编码量表确定了预防犯罪评价系统的5个维度,提出了编码方案并用于评估刑事司法中证据的正确性;李新德等^[4]从内、外证据不确定性分析的角度深入揭示了证据冲突产生的原因,提出了一种同时考虑证据自冲突和外部冲突的相似性测度,然后利用新测度计算证据的众信度,对证据源进行修正;刘艳等^[5]在对网络入侵进行取证的过程中,结合证据推理的思想原理,提出了一种证据推理的弱关联网络入侵的取证方法;田志宏等^[6]基于实时取证的思想,提出了一种支持犯罪重现的按需取证技术——DFR2(on-demand forensic technology support for rollback recovery),DFR2缩小了处理范围,缩短了取证时间,提高了证据关联性。但是,以上研究大多是针对网络犯罪进行的取证考虑,缺少对于冲突证据的处理技巧,对实际犯罪案件缺少针对性的思考。

信息融合是以控制理论、系统理论等基础性理论为基石发展起来的一门综合性理论,随着现代信息处理技术的高速发展,信息融合手段得到越来越广泛的应用。Campomanes-Alvarez等^[7]提出了一个基于模糊聚合函数的信息融合层次框架,以支持决策阶段的人类学专家在不同层次使用该框架来实现颅面解剖的自动化;Xu等^[8]提出了内部置信度和外部置信度来估计多源信息系统中每个信息源的可靠性;何晶晶等^[9]针对现实中同一实体存在不同表象的问题,提出了一种基于证据理论特征融合的同义实体识别方法;张亮等^[10]引入信息融合技术对不同数据挖掘分类方法得到的结果进行融合处理,解决了不同数据挖掘方法得到的结果不一致的问题。但在以往的实际研究中,运用信息融合手段对警务警情及犯罪情报进行处理的研究尚不多见。

针对具有多元冲突信息特征的证据集合,可以采取结合D-S证据推理融合的方法来处理此类问题。Dempster^[11]首先提出了基于人工智能范畴的证据理论,随后Shafer^[12]进一步发展了一种不确定目标的推理理论,与传统的贝叶斯概率论相比,它可以更好地把握问题的无知性和不确定性。在D-S证据融合方法被提出以来,由于其强有力的组合公式能够融合很多来源不同的证据,学者对信息融合的研究日益增多,其在不确定推理和数据融合过程中得到了广泛的应

用。龚本刚等^[13]针对含有投入产出指标的混合型多属性决策问题,提出了一种基于证据理论和数据包络分析(DEA)交叉效率的决策方法;李鹏等^[14]针对方案的指标值为区间直觉模糊数的决策问题,通过D-S合成法则进行信息融合,提出了一种基于灰色关联分析和D-S证据理论的决策方法。

然而,在实际应用研究中,受到外界环境、人为干扰等因素的影响,直接应用D-S理论进行证据融合经常会产生与常理相悖的结论。因此,国内外学者对D-S理论的合成规则进行了更加深入的研究^[15-17]。Jiang等^[15]提出了一种改进的广义证据组合规则,更有效地结合了概率分配函数来组合证据;毛艺帆等^[16]通过取小累加证据体之间两个相交为非空集的焦元BPA,得到证据重合度,然后基于其与证据冲突程度的关系,计算出证据冲突大小;陈圣群等^[17]结合批量式融合和序贯式融合的优点,提出了证据分组合成法,在焦元数目不多的情况下适合大量证据的融合。

为了有效融合犯罪案件中罪犯指证问题中的高冲突证据,本文提出一种新的基于证据可信度和灰色关联的证据融合方法。从改变证据源入手,以灰色相似关联度作为证据的联系来度量证据源中各个证据之间的相似性程度;考虑到证据的可信度,引入证据之间目标的贴适度,用可信度作为权重,对参与融合的证据的基本概率分配函数进行加权平均,再利用Dempster组合规则对这些平均后的证据进行融合。通过实际罪案仿真实验结果比较,表明了本文方法能够较好地处理冲突证据,提高证据组合的可靠性和合理性。

1 基本概念与理论

1.1 证据理论

定义1^[18] 证据理论的一个样本空间称为一个辨识框架,用 Ω 表示。 Ω 是由一系列互不相容的对象组成的有限集合,且包含当前需要识别的全部对象,表示对某一对象的所有可能答案,但其中只有一个正确答案,即

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\},$$

其中对象 ω_i 称为 Ω 的一个单子。只含有一个单子的集合被称为单子集合。

这种单子就是在数据融合系统中,系统要做出决策或判断的结论。证据理论的基本原理是:已知辨识框架 Ω ,判断 Ω 中的一个先验证的未定元素属于 Ω 中某一个子集的程度。对于每一个子集,可以指派一个概率,即基本概率分配,定义如下。

定义2^[18] $P(\Omega)$ 为 Ω 的所有子集构成的集合, $P(\Omega)$ 为 Ω 的幂集, ϕ 为空集,称 $m: P(\Omega) \rightarrow [0, 1]$ 为

基本概率分配函数(BPA,也称之为Mass函数),满足

$$\sum_{A \in P(\Omega)} m(A) = 1, m(\phi) = 0.$$

对于任意的 $A \subseteq \Omega$,若有 $m(A) > 0$,则称 A 为焦元.

定义3^[12](D-S合成规则) 对于 $\forall A \in \Omega, m_1$ 和 m_2 为 Ω 上的两个概率分配函数(Mass函数), B 和 C 为焦元,则D-S合成规则为

$$m_1 \oplus m_2(A) = \frac{1}{K} \sum_{B \cap C = A} m_1(B) \cdot m_2(C). \quad (1)$$

其中 K 为归一化常数,表示为

$$K = \sum_{B \cap C \neq \emptyset} m_1(B) \cdot m_2(C) = 1 - \sum_{B \cap C = \emptyset} m_1(B) \cdot m_2(C),$$

K 的大小反映了证据之间冲突的程度.

1.2 D-S合成规则存在的问题及改进方法

证据理论在某些情况下(如高冲突、证据源少)得出来的结论往往与实际不一样,会出现与实际相悖的现象.因此,学者们给出了不同的改进方法,主要可以从以下两个方面考虑.

1.2.1 第1类改进方法

从改进D-S组合规则的方面考虑.在某种情况,如高冲突证据融合的情况下,直接使用D-S组合规则得到的结果往往是不正确的,且这些错误的结果是在归一化计算的过程中所产生的,所以这一类方法需要从归一化计算过程下手.主要从两个部分考虑:重新分配冲突证据时,被分配到子集的选择;选择了子集后,分配冲突证据的比例.这种情形下,可以采取统一信度函数组合方法,令

$$m(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C) + m^c(A), \\ \forall A, B, C \subseteq \Omega.$$

其中 $m^c(A)$ 表示冲突信息中分配给子集 A 的BPA,也可以写成

$$\begin{cases} m^c(A) = w(A, m) \cdot m(\phi), \forall A \subseteq P; \\ m^c(A) = 0, \text{ else.} \end{cases}$$

P 为重新分配冲突的子集的集合,而 $w(A, m)$ 是权重,且满足

$$\sum_{A \subseteq \Omega} w(A, m) = 1,$$

它决定了分配给各个子集冲突的大小.

1.2.2 第2类改进方法

从改变原始证据源的方面考虑.此类方法将不改变经典Dempster组合规则,而在处理原始证据源的方面进行研究.考虑到证据与证据之间的相关程

度,进行证据融合时,先处理互相冲突的证据,再用D-S组合规则进行融合.即,先对证据 m_1 和 m_2 进行处理,得出新证据 m'_1 和 m'_2 ,之后用D-S组合规则融合证据,即 $(m'_1 \oplus m'_2)$.

源于这一思想,Murphy^[19]提出了基于改变证据源方法的组合规则,称之为平均合成规则.它的主要过程是,平均需要融合的证据的BPA后,利用D-S组合规则对这些平均后的证据进行融合.这一方法虽然改进了不少,但仍有缺陷.因为平均合成法只考虑到了“平均”,只是对证据的简单平均,却没有考虑到证据与证据之间的相互联系.

1.3 序列相似关联度

为度量系统中各个证据间的联系和相似性程度,引入以下序列相似关联度的概念.

定义4^[20] 设系统行为序列 $X_i = (x_i(1), x_i(2), \dots, x_i(n))$,且

$$X_i D = (x_i(1)d, x_i(2)d, \dots, x_i(n)d).$$

其中: D 为序列算子, $x_i(k)d = x_i(k) - x_i(1), k = 1, 2, \dots, n$,则称 D 为零点始化算子; $X_i D$ 为 X_i 的零点始化项,记

$$X_i D = X_i^0 = (x_i^0(1), x_i^0(2), \dots, x_i^0(n)).$$

故 X_i, X_j 的零点始化项分别为

$$X_i^0 = (x_i^0(1), x_i^0(2), \dots, x_i^0(n)), \\ X_j^0 = (x_j^0(1), x_j^0(2), \dots, x_j^0(n)).$$

令

$$s_i - s_j = \int_1^n (X_i^0 - X_j^0) dt \quad (2)$$

是测度序列的零点始化项从1到 n 的积分,不难看出 $s_i - s_j$ 的符号是不定的.称序列 X_i 的各观测数据间的距离之和为 X_i 的长度.

定义5^[20](灰色相似关联度) 设序列 X_i 与 X_j 长度一样,则称

$$\varepsilon_{ij} = \frac{1}{1 + |s_i - s_j|} \quad (3)$$

为 X_i 与 X_j 的基于相似性视角的灰色关联度,简称灰色相似关联度.

2 基于证据可信度和相似关联度的冲突证据融合方法

在对证据体融合的过程中,部分学者曾引入距离函数的概念来衡量距离的相似程度^[21-22],但随着证据体的形式和数量增多,会产生计算量增大、融合困难等问题.灰色相似关联度从相似性视角测度序列之间相互关系和影响,可以用于测度证据序列在几何形状上的相似程度^[20],因而运用相似关联分析可以寻求证据系统中各子证据体之间的数值关系.此外,

相比于其他的证据距离函数,用灰色相似关联度来测度证据体还具有计算简洁、几何意义明显等优势.因此,灰色相似关联度分析对于证据系统发展变化态势提供了量化的度量,适合进行证据体的动态分析.基于此,本文提出一种新的基于灰色相似关联度和冲突证据融合的方法,并将之应用于罪案推理过程.首先给出以下定理.

定理1 令 $\varepsilon_{ij} = \frac{1}{1 + |s_i - s_j|}$ 为长度相同序列 X_i 与 X_j 的灰色相似关联度,则 ε_{ij} 构成的距离测度矩阵 $D = [\varepsilon_{ij}] = [d(s_i, s_j)]$ 必满足以下条件: 1) $d(s_i, s_j) \leq 1$ 且 $d(s_i, s_j) = 1$, 当且仅当 $s_i = s_j$; 2) s_i 和 s_j 越接近,则 $d(s_i, s_j)$ 越接近1; 3) s_i 和 s_j 越相离,则 $d(s_i, s_j)$ 越接近0. 也就是说,测度矩阵 D 反映了序列 X_i 与 X_j 的“逆冲突”状态.

证明 1) $|s_i - s_j| \geq 0$, 则易知 $\frac{1}{1 + |s_i - s_j|} \leq 1$, 当 $s_i = s_j$ 时, $\frac{1}{1 + |s_i - s_j|} = 1$, 反之亦然;
2) 若 $s_i \rightarrow s_j$ 或 $s_j \rightarrow s_i$ 时, 有 $|s_i - s_j| \rightarrow 0$, 则 $\frac{1}{1 + |s_i - s_j|} \rightarrow 1$;
3) 与2)同理. \square

由此可见, $D = [\varepsilon_{ij}]$ 的确定义了一个距离测度矩阵,并且考虑了证据序列 X_i 与 X_j 的相似性属性,故以下考虑冲突证据之间的相似性,利用此距离测度矩阵来构造一个相似性测度矩阵.

定义6 (相似性测度矩阵) 令

$$\text{Com}(m_i, m_j) = C_{ij}, i, j = 1, 2, \dots, n$$

为Mass函数 m_i 与 m_j 的相似性测度,也可以用矩阵来表示,即它的相似测度矩阵为

$$\text{COM} = \begin{bmatrix} 1 & C_{12} & C_{13} & \dots & C_{1n} \\ C_{21} & 1 & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \dots & \vdots \\ C_{n1} & C_{n2} & C_{n3} & \dots & 1 \end{bmatrix}, \quad (4)$$

其中测度矩阵的主对角元素为1.相似性测度矩阵表明,两个证据体在几何形状上越相似,它们的相似性测度越大.

定义7 令 $\text{Com}(m_i, m_j) = C_{ij}$ 为Mass函数 m_i 与 m_j 的相似性测度,则

$$\text{Sup}(m_i) = \sum_{j=1, j \neq i}^n \text{Com}(m_i, m_j) = \sum_{j=1, j \neq i}^n C_{ij} \quad (5)$$

为 m_i 的支持度, $C_{ij} \in \text{COM}, i, j = 1, 2, \dots, n$. 令

$$\text{Cfl}(m_i) = \frac{\text{Sup}(m_i)}{\sum_{j=1}^n \text{Sup}(m_j)}, i, j = 1, 2, \dots, n \quad (6)$$

为系统中证据 m_i 的可信度,且满足

$$\sum_{i=1}^n \text{Cfl}(m_i) = 1.$$

由定义7可知,可信度 $\text{Cfl}(m_i)$ 是将相似测度矩阵中每一行除自身相似测度之外的所有元素求和,证据体Mass函数 m_i 的支持度 $\text{Sup}(m_i)$ 反映的是 m_i 被其他证据所支持的程度,可见 $\text{Cfl}(m_i)$ 是相似性测度 C_{ij} 的函数.如果一个证据体与其他证据体相似性测度比较高,则认为它们相互支持的程度也高,这些证据相互支持对方;如果一个证据与其他证据相似性测度较低,则认为它们相互支持的程度也低.将支持度归一化后得到可信度,可信度反映的是一个证据的可信程度.由此分析可知,一个证据被其他证据所支持的程度越高,该证据越可信.如果一个证据不被其他证据所支持,则认为该证据的可信度较低.在这种情形下,一个证据被其他证据所支持的程度多少是由可信度来衡量的.

定义8 称

$$m_{\text{AE}}(X) = \sum_{i=1}^n (\text{Cfl}(m_i) \times m_i(X)) \quad (7)$$

为证据集合 X 基于可信度的加权平均证据,其中 $X \in \{A | A \subseteq P(\Omega), \Omega \text{ 为已知辨识框架}\}$.

基于以上,提出基于证据可信度和灰色相似关联度的冲突罪案证据融合方法.新的证据融合方法是从改变罪案证据源开始入手,用证据可信度作为权重,对参与融合的罪案证据的BPA,即Mass函数进行加权平均,得到新的减弱冲突的罪案证据,最后利用D-S组合规则对平均后的罪案证据进行融合.具体步骤如下.

Step 1: 系统收集到 n 个罪案证据 m_1, m_2, \dots, m_n , 由定理1分别计算这些罪案证据的灰色相似关联度 $\varepsilon_{ij}(i, j = 1, 2, \dots, n)$, 以 ε_{ij} 为元素得到距离测度矩阵 $D = [\varepsilon_{ij}]_{n \times n}$;

Step 2: 从证据相似度出发,由定义7将罪案证据相似测度矩阵COM用距离测度矩阵 D 代替,得到 $\text{Com}(m_i, m_j) = C_{ij}, i, j = 1, 2, \dots, n$, 计算罪案证据 m_i 的支持度 $\text{Sup}(m_i)$, 并将支持度归一化后得到罪案证据 m_i 的可信度 $\text{Cfl}(m_i)$;

Step 3: 将证据可信度 $\text{Cfl}(m_i)$ 作为权重,由定义8对参加融合的罪案证据BPA进行加权平均,即 $m_{\text{AE}}(X) = \sum_{i=1}^n (\text{Cfl}(m_i) \times m_i(X))$, 得到减弱冲突影响的新证据 m'_1, m'_2, \dots, m'_n ;

Step 4: 由定义3,使用D-S证据合成规则融合这些加权平均后的新罪案证据 m'_1, m'_2, \dots, m'_n , 由得到的概率分配函数BPA对罪案证据进行判断.

3 案例分析

罪案的诉讼活动实际上是围绕证据的收集、分析、判断、质证和采纳而进行的,证据问题是罪案诉讼中的核心问题,而证据的厘定离不开证据标准的确定.《中华人民共和国刑事诉讼法》规定的证明标准明确描述了证据标准的内涵,该法第195条规定法院判定被告人有罪的证明标准是“案件事实清楚,证据确实、充分”,第53条第2款规定证据确实、充分,应当符合以下条件:1)定罪量刑的事实都有证据证明;2)据以定案的证据均经法定程序查证属实;3)综合全案证据,对所认定事实已排除合理怀疑.此外,《中华人民共和国刑事诉讼法》第五章第四十八条分别从内涵和外延角度对证据进行了界定:可以用于证明案件事实的材料都是证据.证据包括:1)物证;2)书证;3)证人证言;4)被害人陈述;5)犯罪嫌疑人、被告人供述和辩解;6)鉴定意见;7)勘验、检查、辨认、侦查实验等笔录;8)视听资料、电子数据.

罪案推理的过程实际上是侦查假说的形成和证实过程.无论是侦查假说的形成还是侦查假说的证实,都在很大程度上需要借助于证据的推理.罪案侦查中运用的推理称为侦查推理,它是形成侦查假说的主要方式.用以证实或者检验侦查假说的罪案证据称为侦查证据,它是侦查假说证实或者检验中的核心问题,侦查证据因案情事实的不同而呈现出不同的类型,它的采用需要遵守一些逻辑和法律的标准.在罪案侦查中,办案人员认知案情事实有两种方法:一种是所谓的感知方法,即借助于感官的经验认知方法;一种是推知方法,即借助于大脑的思维认知方法^[23].其中:借助于大脑的思维认知方法主要是推断,而推断是一个很复杂的问题,办案人员经常能够作出推断,但不一定能经常解释推断,这是实际办案过程中经常面对的困难.

在罪案推理过程中,认证证据力要以证据与案件事实的相关性为依据,很多情况下,由于证据提供者为主观性和证据内容的客观真实性无法百分之百地确定,而只能推断其可能性的大小,或者以某种分布的形式来给出证据指向,可见不确定性的证据在罪案推理问题中是广泛存在的.在大数据时代,依靠经验认知方法来处理不确定性的犯罪冲突证据显然是不适用的,而借助于大脑的思维认知方法,通过建立数学模型来处理犯罪冲突证据是新的可尝试路径,在目前来看尚不多见.

以安徽省某市公安局2016年侦破的某区入室盗窃类刑事犯罪串并案为例,采用基于信息融合的灰色关联证据融合方法来处理案件中的冲突证据.自

2014年12月到2015年12月,该市某区发生多起入室盗窃案件,被盗财物包括现金、电脑、手机、烟酒、首饰等易于携带物品,损失价值达到50余万元,此系列案件犯罪分子在作案时间、作案手段、侵害物品以及现场遗留痕迹物证上具有一定程度的相似性.该市刑警大队同该区所属辖派出所通过现场勘验、调查走访、视频追踪、警务信息平台查询等方法,锁定了犯罪嫌疑人A、B、C三人(即辨识框架 $\Omega = \{A, B, C\}$).该刑事犯罪串并案提供的证据具有典型的相似性和不确定性特征,在对证人证词及其他证据的综合考虑基础上,对A、B、C三人分别指证,得到5个侦查证据结果 m_1, m_2, m_3, m_4, m_5 (即BPA).其中: m_1 为证人证言, m_2 为被害人陈述, m_3 为警局鉴定意见, m_4 为犯罪嫌疑人的供述辩解, m_5 为案发当时周边的电子视频监控数据.具体如表1所示.

表1 证据对嫌疑人的概率指证结果

嫌疑人	m_1	m_2	m_3	m_4	m_5
A	0.5	0	0.55	0.55	0.55
B	0.2	0.9	0.1	0.1	0.1
C	0.3	0.1	0.35	0.35	0.35

根据收集到的5个侦查证据 $m_1 \sim m_5$,由式(3)分别计算这些证据的灰色相似关联度 $\varepsilon_{ij}(i, j = 1, 2, \dots, 5)$,以 ε_{ij} 为元素得到距离测度矩阵D;从证据相似度角度出发,由式(4)可得证据相似测度矩阵为

$$\text{COM} = [\text{Com}(m_i, m_j)]_{5 \times 5} =$$

$$\begin{bmatrix} 1 & 0.4255 & 0.8696 & 0.8696 & 0.8696 \\ 0.4255 & 1 & 0.4000 & 0.4000 & 0.4000 \\ 0.8696 & 0.4000 & 1 & 1 & 1 \\ 0.8696 & 0.4000 & 1 & 1 & 1 \\ 0.8696 & 0.4000 & 1 & 1 & 1 \end{bmatrix}.$$

由式(5)可得证据的支持度分别为 $\text{Sup}(m_1) = 3.0343$, $\text{Sup}(m_2) = 1.6255$, $\text{Sup}(m_3) = 2.8255$, $\text{Sup}(m_4) = 3.2696$, $\text{Sup}(m_5) = 3.2696$.

由式(6)可得证据 m_i 的可信度分别为 $\text{Cfl}(m_1) = 0.2163$, $\text{Cfl}(m_2) = 0.1159$, $\text{Cfl}(m_3) = 0.2015$, $\text{Cfl}(m_4) = 0.2331$, $\text{Cfl}(m_5) = 0.2331$.

将可信度 $\text{Cfl}(m_i)$ 作为权重,由式(7)对参加融合的证据BPA进行加权平均,得到 $m_{\text{AE}}(A) = 0.4754$, $m_{\text{AE}}(B) = 0.2144$, $m_{\text{AE}}(C) = 0.3102$.由此得到减弱冲突影响的新犯罪证据 $m'_i, i = 1, 2, \dots, 5$.使用D-S证据合成规则融合这些加权平均后新证据 m'_1, m'_2, \dots, m'_5 ,最后得到概率分配函数BPA,从而判断具有最大可能的是犯罪嫌疑人A.

为了使分析结果具有可比性,分别采用文献[12, 17, 19, 24, 25]的方法对所给案例进行分析,所得对比结果如表2所示.

表2 罪案推理方法证据融合结果对比

融合规则	m_1, m_2, m_3	m_1, m_2, m_3, m_4	m_1, m_2, m_3, m_4, m_5	决策结果
D-S ^[11]	$m(A) = 0$ $m(B) = 0.6316$ $m(C) = 0.3684$	$m(A) = 0$ $m(B) = 0.3288$ $m(C) = 0.6712$	$m(A) = 0$ $m(B) = 0.1228$ $m(C) = 0.8772$	C
Yager ^[24]	$m(A) = 0$ $m(B) = 0.018$ $m(C) = 0.0105$ $m(\Omega) = 0.9715$	$m(A) = 0$ $m(B) = 0.0018$ $m(C) = 1.00368$ $m(\Omega) = 0.99452$	$m(A) = 0$ $m(B) = 0.00018$ $m(C) = 0.00129$ $m(\Omega) = 0.99853$	Ω
文献[19]	$m(A) = 0.3500$ $m(B) = 0.5224$ $m(C) = 0.1276$	$m(A) = 0.6027$ $m(B) = 0.2627$ $m(C) = 0.1346$	$m(A) = 0.7958$ $m(B) = 0.0932$ $m(C) = 0.1110$	A
文献[25]	$m(A) = 0.5018$ $m(B) = 0.1481$ $m(C) = 0.2652$ $m(\Omega) = 0.0849$	$m(A) = 0.6889$ $m(B) = 0.0622$ $m(C) = 0.2442$ $m(\Omega) = 0.0048$	$m(A) = 0.7942$ $m(B) = 0.0180$ $m(C) = 0.1870$ $m(\Omega) = 0.0008$	A
文献[17]	$m(A) = 0.6289$ $m(B) = 0.1013$ $m(C) = 0.2698$	$m(A) = 0.7444$ $m(B) = 0.0358$ $m(C) = 0.2198$	$m(A) = 0.8213$ $m(B) = 0.0118$ $m(C) = 0.1669$	A
本文方法	$m(A) = 0.5249$ $m(B) = 0.3098$ $m(C) = 0.1652$	$m(A) = 0.7700$ $m(B) = 0.0771$ $m(C) = 0.1528$	$m(A) = 0.8796$ $m(B) = 0.0164$ $m(C) = 0.1040$	A

从表2可以看出: D-S融合规则和Yager融合规则对嫌疑人A的融合结果始终是0, 显然与常理相悖. 由于D-S融合规则自身的缺陷, 不适合对冲突证据进行融合. Yager规则虽然在D-S规则的基础上进行了修正, 但其本身忽略了冲突信息的可用性, 夸大了证据冲突的程度, 不能有效处理冲突证据, 因此两者都导致了错误的罪案判断结果. Murphy^[19]的平均法方法虽然可以推断出犯罪嫌疑人A, 但从表2的数据不难看到, 该方法在收集到第5个证据的时候, 才能比较清晰地识别出犯罪嫌疑人A, 且其计算量较大, 用到的计算方法相对复杂. 文献[25]和文献[17]分别利用参考证据对各原始证据进行修正和基于分组合成的思想, 取得了比经典方法满意的效果, 基本在收集到第4个证据时, 就能快速识别到犯罪嫌疑人A, 但所用方法较为繁琐. 本文方法充分考虑到证据之间的关联性和冲突性, 运用灰色关联思想和证据支持的方法, 虽然在收集到第3个证据时, 犯罪嫌疑人A的分配概率为 $m(A) = 0.5249$, 效果略优于文献[25]的 $m(A) = 0.5018$, 但不如文献[17]的 $m(A) = 0.6289$. 但随着重复证据的出现, 目标指向快速收敛到最终犯罪嫌疑人A, 在最终的识别目标效果中, A的分配概率提高到 $m(A) = 0.8796$, 体现了灰色关联测度的强分辨效果, 且改进的关联测度算法的复杂度相对较低.

4 结论

在实际罪案证据的信息融合环境中, 由于受到人为或自然环境等很多因素的影响, 警务系统收集到的

证据往往存在较大的冲突, 这时如何有效处理及利用这些冲突证据极为重要. 本文在D-S证据理论以及已有修正规则的基础上, 提出了一种结合灰色关联测度的加权证据合成方法, 该方法可以有效地处理冲突证据, 且收敛速度较快, 使得高冲突证据融合结果更加可靠合理. 实际罪案仿真实验的结果比较表明, 本文方法能够较好地处理冲突证据, 提高证据组合的可靠性和合理性, 为社会公共安全领域的实际问题处理提供了参考和借鉴.

参考文献(References)

- [1] 封利强. 我国刑事证据推理模式的转型: 从日常思维到精密论证[J]. 中国法学, 2016(6): 153-176. (Feng L Q. The transformation of criminal evidence inference pattern in China: From daily thinking to sophisticated argument[J]. China Legal Science, 2016(6): 153-176.)
- [2] Kennedy L W, Caplan J M, Piza E L, et al. Vulnerability and exposure to crime: Applying risk terrain modeling to the study of assault in Chicago[J]. Applied Spatial Analysis and Policy, 2016, 9(4): 529-548.
- [3] Johnson S D, Tilley N, Bowers K J. Introducing EMMIE: An evidence rating scale to encourage mixed-method crime prevention synthesis reviews[J]. J of Experimental Criminology, 2015, 11(3): 459-473.
- [4] 李新德, 王丰羽. 一种基于ISODATA聚类和改进相似度的证据推理方法[J]. 自动化学报, 2015, 41(3): 575-590. (Li X D, Wang F Y. A method of evidence reasoning based on ISODATA clustering and improved similarity measure[J]. Acta Automatica Sinica, 2015, 41(3): 575-590.)

- [5] 刘艳, 丁颖. 弱关联网络入侵的取证技术仿真分析[J]. 计算机仿真, 2016, 33(4): 340-343.
(Liu Y, Ding Y. Simulation analysis of forensics technology for weak link network intrusion[J]. Computer Simulation, 2016, 33(4): 340-343.)
- [6] 田志宏, 姜伟, 张宏莉. 一种支持犯罪重现的按需取证技术[J]. 清华大学学报: 自然科学版, 2014, 54(1): 20-28.
(Tian Z H, Jiang W, Zhang H L. On-demand forensics to support crime scene reconstruction[J]. J of Tsinghua University: Science and Technology, 2014, 54(1): 20-28.)
- [7] Campomanes-Alvarez C, Ibáñez O, Cordón O, et al. Hierarchical information fusion for decision making in craniofacial superimposition[J]. Information Fusion, 2018, 39: 25-40.
- [8] Xu W, Yu J. A novel approach to information fusion in multi-source datasets: A granular computing viewpoint[J]. Information Sciences, 2017, 378: 410-423.
- [9] 何晶晶, 蔡德胜, 介飞, 等. 利用D-S证据理论进行特征融合的同义实体识别[J]. 计算机应用研究, 2018(5): 1-6.
(He J J, Cai D S, Jie F, et al. Synonymous entity recognition based on d-s evidence theory for feature fusion[J]. Application Research of Computers, 2018(5): 1-6.)
- [10] 张亮, 张玲玲, 陈懿冰, 等. 基于信息融合的数据挖掘方法在公司财务预警中的应用[J]. 中国管理科学, 2015, 23(10): 170-176.
(Zhang L, Zhang L L, Chen Y B, et al. Based on information fusion technique with data mining in the application of finance early-warning[J]. Chinese J of Management Science, 2015, 23(10): 170-176.)
- [11] Dempster A P. Upper and lower probabilities induced by a multivalued mapping[J]. Annual Mathematics and Statistics, 1967, 38(2): 325-339.
- [12] Shafer G A. Mathematical theory of evidence[M]. Princeton: Princeton University Press, 1976: 58-62.
- [13] 龚本刚, 张孝琪, 郭丹丹. 基于证据理论-DEA交叉效率的混合型多属性决策方法[J]. 控制与决策, 2016, 31(5): 943-948.
(Gong B G, Zhang X Q, Guo D D. Method for hybrid multiple attribute decision-making based on Dempster-Shafer theory and cross efficiency of DEA[J]. Control and Decision, 2016, 31(5): 943-948.)
- [14] 李鹏, 刘思峰. 基于灰色关联分析和D-S证据理论的区间直觉模糊决策方法[J]. 自动化学报, 2011, 37(8): 993-998.
(Li P, Liu S F. Interval-valued intuitionistic fuzzy numbers decision-making method based on grey incidence analysis and D-S theory of evidence[J]. Acta Automatica Sinica, 2011, 37(8): 993-998.)
- [15] Jiang W, Zhan J. A modified combination rule in generalized evidence theory[J]. Applied Intelligence, 2017, 46(3): 630-640.
- [16] 毛艺帆, 张多林, 王路. 基于重合度的证据冲突度量方法[J]. 控制与决策, 2017, 32(2): 293-298.
(Mao Y F, Zhang D L, Wang L. Measurement of evidence conflict based on overlapping degree[J]. Control and Decision, 2017, 32(2): 293-298.)
- [17] 陈圣群, 王应明. 证据的分组合成法[J]. 控制与决策, 2013, 28(4): 574-578.
(Chen S Q, Wang Y M. Grouping method for combining evidence[J]. Control and Decision, 2013, 28(4): 574-578.)
- [18] 康耀红. 数据融合理论与应用[M]. 西安: 西安电子科技大学出版社, 2006: 141-148.
(Kang Y H. The Theory of data fusion and its application[M]. Xi'an: Xidian University Press, 2006: 141-148.)
- [19] Murphy C K. Combining belief functions when evidence conflicts[J]. Decision Support Systems, 2000, 29(1): 1-9.
- [20] 刘思峰, 杨英杰, 吴利丰, 等. 灰色系统理论及其应用[M]. 第7版. 北京: 科学出版社, 2014: 66-90.
(Liu S F, Yang Y J, Wu L F, et al. Grey system theory and its application[M]. 7th ed. Beijing: Science Press, 2014: 66-90.)
- [21] Deng Y, Shi W K, Zhu Z F. Efficient combination approach of conflict evidence[J]. J of Infrared and Millimeter Waves, 2004, 23(1): 27-32.
- [22] Jousselme A L, Grenier D, Éloi Bossé. A new distance between two bodies of evidence[J]. Information Fusion, 2001, 2(2): 91-101.
- [23] 马前进. 公安机关刑事案件侦查中的假说、推理和证据[J]. 广州市公安管理干部学院学报, 2016, 26(2): 40-45.
(Ma Q J. Hypothesis, reasoning and evidence in the investigation of criminal cases in public security organs[J]. J of Guangzhou Police College, 2016, 26(2): 40-45.)
- [24] Yager R R. On the Dempster-Shafer framework and new combination rules[J]. Information System, 1989, 41(2): 93-137.
- [25] 熊彦铭, 杨战平, 屈新芬. 基于模型修正的冲突证据组合新方法[J]. 控制与决策, 2011, 26(6): 883-887.
(Xiong Y M, Yang Z P, Qu X F. Novel combination method of conflict evidence based on evidential model modification[J]. Control and Decision, 2011, 26(6): 883-887.)

作者简介

徐祖润(1979—), 男, 讲师, 博士生, 从事灰色系统理论和复杂系统建模的研究, E-mail: xuzurun@163.com;

刘思峰(1955—), 男, 教授, 博士生导师, 从事灰色系统理论、系统方法与模型等研究, E-mail: sfliu@nuaa.edu.cn;

方志耕(1962—), 男, 教授, 博士生导师, 从事不确定性理论与质量管理等研究, E-mail: zhigengfang@163.com;

吴洁(1968—), 女, 教授, 博士生导师, 从事技术创新及管理、知识产权保护与管理等研究, E-mail: 0511wujie@163.com;

刘勇(1985—), 男, 副教授, 博士, 从事科技管理、冲突分析等研究, E-mail: cly1985528@163.com.

(责任编辑: 齐 霖)