

水中目标信息智能识别方法研究

王冠, 郭育

(中国船舶工业系统工程研究院, 北京 100094)

摘要 提出了基于判定准则的海洋环境下目标关系识别方法。通过 HBACA 算法查找或识别不同元素关系能力强的特性,提取目标信息数据的关系,之后在信息数据关系中计算每个信息节点的点度中心度、紧致中心度和节点活跃度,根据信息数据在目标事件中的重要或关键程度,将信息数据划分为关键目标信息、重要目标信息和一般目标信息。仿真结果显示:能够明确地获得关键的信息数据坐标,并与其所处事件的重要程度能够匹配,同时相关信息数据的关系非常明显,很容易提炼出所处目标事件中重要目标信息数据之间的关联。

关键词 水中目标;目标信息;识别方法;智能算法

中图分类号 E917 **文献标识码** A

Research on Intelligent Recognition Method of Underwater Target Information

WANG Guan, GUO Yu

(China Shipbuilding Industry System Engineering Research Institute, Beijing 100094, China)

Abstract A target recognition method under ocean environment based on decision criterion is proposed. Through the HBACA algorithm, this paper finds or identifies the characteristics of the strong relationship between different elements, extracts the relationship between target information data, and then calculates the degree centrality, compact centrality and node activity of each information node in the information data relationship. According to the important or critical degree of information data in the target event, the information data is divided into key information, important information and general information. The simulation results show that the key information data coordinate can be obtained clearly, and it can match the important degree of the event. At the same time, the relation of the related information data is very obvious, and it is easy to extract the association between the important information data in the target event.

Key words underwater target; target information; recognition method; intelligent algorithm

0 引言

随着海洋基础设施建设的步伐向前迈进,海洋中的目标信息数据量与日俱增已成为现实趋势,进行目标信息数据识别,将不同海洋目标信息按照重要和关键程度进行关系划分,并赋予每条目标信息数据一种关系^[1]。目前有学者采用经典的 HBA-

CA 算法^[2],对大数据的每条信息的重要性进行大致定义,但他们的研究成果,在区分信息目标数据的重要性问题上有待优化与完善。

本文首先利用 HBACA 算法查找或识别不同元素关系能力强的特性,提取目标信息数据的关系,主要是通过搜索蚁的行为特性中混入侦察蚁的行为特性,即搜索蚁在搜索过程中,会找到侦查

蚁已经标记的信息素进行重复标记,将重复标记的信息素链路作为信息数据的关系网,进而识别出大数据中信息数据的关系。然后,在信息数据关系中计算每个信息节点的点度中心度、紧致中心度和节点活跃度,根据信息数据在目标事件中的重要或关键程度,将信息数据划分为关键目标信息、重要目标信息和一般目标信息,提出了信息数据的判定准则,形成了基于关系的信息数据识别方法。

1 目标分类

水中目标分类在军事防御及进攻上具有重要作用,目标特征与海洋信道之间的关联十分紧密,水声探测系统接收到的目标信号都是目标源激发信号与信道多途响应耦合后的信号,本文基于声呐已探测目标信息内容,从水面目标和水下目标所激发的声场波数谱差异特征入手,将目标分为 3 类。

一般目标(节点):浮漂、侦察船或渔船、邮轮、集装箱船、货轮、LNG 船、海警船。

重要目标(节点):驱逐舰、护卫舰、常规潜艇、UUV、特种潜航器、蛙人。

关键目标(节点):航母、战略潜艇、鱼雷、水雷,战略趸船。

2 信息数据识别

HBACA 算法,即混合蚂蚁算法,可以识别出网络数据中信息节点之间的关系,它是基于蚁群算法进行的改进,其主要思想是:通过引入具有多行为的混合蚂蚁来扩大搜索解空间,避免局部最优解、早熟和停滞现象;另外在每次迭代过程中具有不同行为的蚂蚁数目可以视具体情况而进行动态调整,这样就可以在加速收敛和防止早熟、停滞显现之间取得较好的平衡^[3]。

混合蚂蚁是一种具有多行为的蚂蚁,它包括以下 3 种类型的蚂蚁行为。

行为 1:蚂蚁具有信息感知能力,搜索前进策略在尚未走过的路径上选择信息量大的路径前进,在走完全部路径后对所走路径进行一次全局信息更新,此蚂蚁称为智能蚂蚁。

行为 2:蚂蚁不具有信息感知能力,将其随机

选择一个本次遍历中尚未走过的路径前进,以保证算法搜索更大的解空间,此蚂蚁称为随机蚂蚁。

行为 3:具有与智能蚂蚁相反的搜索前进策略,即选取下一个节点的转移概率与智能蚂蚁正好相反,而其他特征与智能蚂蚁完全相同,此蚂蚁称为逆反蚂蚁。

混合蚁群算法在搜索路径问题上,具有并行侦察和搜索的行为特性,并对信息量较大的路径具有良好的正反馈性和较强的鲁棒性,因此其在信息数据关系识别领域具有较好的研究和应用价值。

目标信息数据识别方法主要过程分为 2 个步骤,方法步骤如下。

步骤 1:针对目标信息数据开展聚类,而后设置目标信息数据各种信息标签的预设值,根据具体情况,对目标信息数据的类型、频率、置信度、代价等信息进行综合处理,通过聚类方法计算,分析信息数据为类型,对目标信息数据极其类似的信息数据开展聚类。

步骤 2:完成目标信息数据聚类分析后,采用基于关系的目标信息数据识别方法对单位时间内目标事件进行分析。首先,利用 HBACA 算法提取目标信息数据的关系;然后,通过目标信息数据的判定准则,识别出目标事件中的关键目标信息、重要目标信息和一般目标信息。

3 信息数据关系提取

本文提出的方法,通过调整混合蚂蚁的投放方式,以及其侦察、搜索方式,从而延长信息素的有效时间,进而在全局范围内提高识别准确率,并实现提取信息数据在其目标事件中的关系。

原有 HBACA 算法仅仅是对执行侦察任务的混合蚂蚁数量,与执行搜索、侦察任务的混合蚂蚁数量进行了调配,但并未考虑信息素的持久性与数据流通过的路径之间的关系,以及信息素的持久性与时间的指数关系。

基于关系的目标信息数据识别方法,采用侦察混合蚂蚁、搜索混合蚂蚁开展研究分析。基于建立目标事件数据流向量表的前提,设置侦察混合蚂蚁的行为偏向侦察蚁,搜索混合蚂蚁的行为是侦察蚁

与搜索蚁采用行为的融合,采用侦察混合蚂蚁与搜索混合蚂蚁的不同特征,针对源信息数据之间的数据流开展局部、全局分析,获得节点上信息数据之间的关系。

HBACA 算法基于蚂蚁有不同的行为特性,设定不同行为蚂蚁的比例参数,HBACA 避免了 ACO、ACS 及 AS 显现出的停滞现象,更重要的是 HBACA 中搜索蚁搜索最优解的性能较佳。实验证明 HBACA 较 AOC 系列算法计算效能有很大的提升,相关计算性能表现良好。在利用 HBACA 提取目标事件中信息节点之间的关系时,设定侦察混合蚂蚁对每一个节点作为开展局部查询与搜索的基点。同时使用特殊的信息素标记获得的数据,使用特殊的信息素标记获得的数据目的是,使得搜索混合蚂蚁抵达被标记的数据点后,搜索下一个被标记的数据点时,发挥有效的信息导引。搜索混合蚂蚁凭借侦察混合蚂蚁对每一个节点使用特殊的信息素的浓度及导引信息,针对局域数据节点开展搜索。

3.1 信息素标记与持久性分析

在 HBACA 算法中,可以采用侦察混合蚂蚁进行信息素的标记。将 m 只侦察混合蚂蚁与 m 个节点一一对应进行放置,每只侦察混合蚂蚁以自身所在节点为基点,侦察与自身所在节点周围相连的 $m-1$ 个节点,并将侦察结果结合现有的信息标记信息素,记为 $\text{Flow}[i][j]$,标记 i 节点到 j 节点路径。其中 $\text{Flow}[i][j]$ ($i, j=0, 1, 2, 3, \dots, m-1; i \neq j$) 公式见式(1)。

$$\text{Flow}[i][j] = \begin{cases} \widetilde{\text{Inf}}_{ij} / \text{Inf}_{ij} & \text{节点 } j \text{ 在节点 } i \text{ 的 3 跳以内} \\ 0 & \text{其他} \end{cases} \quad (1)$$

式中: $\widetilde{\text{Inf}}_{ij}$ 是基于节点为中心,至 $m-1$ 个节点中信息流量的最大值。

基于 $\text{Flow}[i][j]$ 的值为基点,设置初始时刻每条路径上的信息量如式(2)所示。

$$T_{ij}(0) = \begin{cases} C \cdot \text{Flow}[i][j] & \text{若 } \text{Flow}[i][j] \neq 0 \\ \frac{C \cdot \widetilde{\text{Inf}}_{ij}}{\text{Inf}_{ij}} & \text{否则} \end{cases} \quad (2)$$

式中: $\widetilde{\text{Inf}}_{ij}$ 是基于 i 节点为中心,至个节点的信息流量的最小值; C 表示初始时刻每条路径上的信息量。

HBACA 算法仅仅是对执行侦察任务的混合蚂蚁数量,与执行搜索、侦察任务的混合蚂蚁数量进行了调配,并未考虑信息素的持久性与数据流通过的路径之间的关系,以及信息素的持久性与时间的指数关系。事实上信息素的持久性与数据流通过路径上的侦察混合蚂蚁标记次数存在求和关系,同时还与时间存在指数关系。信息素与数据流链路之间的关系如图 1 和图 2 所示。

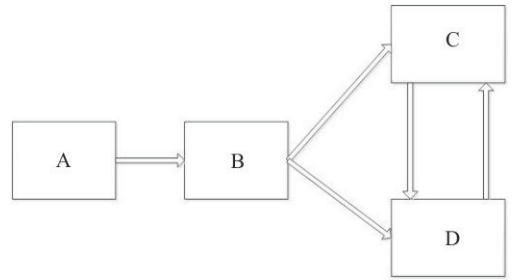


图 1 信息数据节点信息流向图

Fig. 1 Information flow diagram of information data node

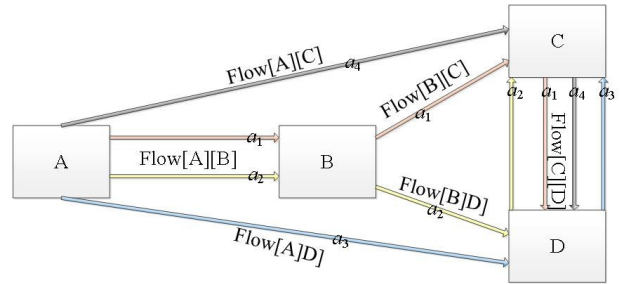


图 2 信息数据节点侦查蚁侦查路径与信息素关系

Fig. 2 Relations between detection paths and pheromones of investigation ant in information data node

如图 1 所示,该图描述了 4 个信息数据节点的信息流向图,分别代表 4 个信息数据节点。图 2 描述了侦查蚁侦查路径与信息素关系,4 只侦察蚁分别为 a_1, a_2, a_3, a_4 , a_1 经过侦察路径为 $A \rightarrow B \rightarrow C \rightarrow D$, a_2 经过侦察路径为 $A \rightarrow B \rightarrow D \rightarrow C$, a_3 经过的侦察路径为 $A \rightarrow D \rightarrow C$, a_4 经过侦察路径为 $A \rightarrow C \rightarrow D$ 。则在计算 D 与 C 两个信息数据节点之间数据流路径上信息素的持久性时,应该采用 4 倍的 $\text{Flow}[C][D]$ 进行计算。在计算 A 与 B 信息数据节点之间数据流路径上信息素的持久性时,采用 2

倍的 $\text{Flow}[A][B]$ 进行计算;同样在计算 B 与 C、B 与 D 信息数据节点之间数据流路径上信息素的持久性时,分别采用 2 倍的 $\text{Flow}[B][C]$ 、 $\text{Flow}[B][D]$ 分别进行计算。因此,如果 $\text{Flow}[A][B] = \text{Flow}[B][C] = \text{Flow}[B][D]$,则由于 D 与 C 两个信息数据节点之间数据链路上信息素的持久性强于 B 与 C 节点、B 与 D 节点数据链路之间信息素持久性。

在侦查混合蚂蚁标记的信息素,会随时间的流

$$P_{ij}^K = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}(t)]^\beta}{\sum_{\text{Flow} \neq \text{tabu}_k} [\tau_{is}(t)]^\alpha [\eta_{is}(t)]^\beta}, & \text{若 } j \notin \text{tabu}_k \text{ 且 } \text{Flow}[i][j] \neq 0 \\ 0 & \text{否则} \end{cases} \quad (3)$$

全部的侦察混合蚂蚁、搜索混合蚂蚁结束一次循环,每条路径上信息量调整为公式(4)所示。

$$\tau_{ij}(t+1) = \begin{cases} \rho \cdot \tau_{ij}(t) + (1-\rho) \cdot \Delta\tau_{ij} & \text{Flow}[i][j] \neq 0 \\ \rho \cdot \tau_{ij}(t) & \text{否则} \end{cases} \quad (4)$$

式中: $\Delta\tau_{ij}$ 是此次循环内全部的侦察混合蚂蚁、搜

$$\rho = \begin{cases} \left[\frac{D_Tran_{ij}^k(t) + D_Disp_{ij}^k(t)}{\sum_{k=1}^n (D_Tran_{ij}^k(t) + D_Disp_{ij}^k(t))} \right]^\varepsilon \cdot (1-\gamma)^\delta \cdot \varphi & 1 < \delta \leq 6 \\ 0 & \delta \geq 6 \end{cases} \quad (5)$$

式中: $D_Tran_{ij}^k(t)$ 、 $D_Disp_{ij}^k(t)$ 是由 i 节点到 j 节点的数据运输、计算延时; γ 是信息调整常量; δ 是至目标节点的跳数; ε 、 φ 是延时与运输距离的调整系数; ε 、 φ 基于节点单位时间 t 内所经过数据包的信息量而确定。

公式(5)中的 ε 如公式(6)所示。

$$\varepsilon = T_{ij}^{\text{data}}(t) / (R_j^{\text{data}}(t) + T_{ij}^{\text{data}}(t)) \quad (6)$$

式中: $T_{ij}^{\text{data}}(t)$ 是需传输的数据包个数, $R_j^{\text{data}}(t)$ 是需重传数据报个数的回馈信息。

$$\Delta\tau_{ij}^k(t) = \begin{cases} \frac{Q \cdot (\widetilde{\text{Inf}}_{ij} / \text{Inf}_{ij})}{L_k}, & \text{若蚂蚁 } k \text{ 经过 } (i, j) \text{ 且 } \text{Flow}[i][j] \neq 0 \\ 0 & \text{否则} \end{cases} \quad (9)$$

从公式(9)可知,每只搜索混合蚂蚁仅仅在具备上下级关系节点处,由最优解组成的路径上留下适量的信息素;根据公式(7),每只搜索混合蚂蚁仅仅在可能是最优解组的路径上追加信息量。

基于每只搜索混合蚂蚁得到的最优解,也就是信息素最多的路线,结合数据流特征,能够得到 i

逝而慢慢变少,直到消失。为了能通过信息数据节点之间信息流关系的分析,识别信息数据关系,需要强化侦查混合蚂蚁标记的信息素,可以采用主要具有搜索特性的搜索混合蚂蚁来加强信息素。

3.2 基于信息素的信息数据关系网络形成

搜索混合蚂蚁, $K(K=1, 2, \dots, n)$ 在时刻 t 时,由 i 节点至 j 节点的转移概率 P_{ij}^K 可以表示为公式(3)。

搜索混合蚂蚁在路径 (i, j) 上的全部信息量之和;由于信息素的持久性 ρ 与数据流通过的路径有关,同时信息素的持久性 ρ 与时间间隔之间存在指数关系,所以 ρ 为公式(5)所示。

公式(5)中 φ 如公式(7)所示。

$$\varphi = \frac{S_{ij}(t) / (S_{ij}(t) + F_{ij}(t))}{S_{ij}(t-1) / (S_{ij}(t-1) + F_{ij}(t-1))} \quad (7)$$

式中: $S_{ij}(t)$ 是数据流由 i 至 j 路径上经过的次数, $F_{ij}(t)$ 为数据流未在 i 至 j 路径上的次数。

$$\Delta\tau_{ij}^k(t) = \sum_{k=1}^n \Delta\tau_{ij}^k(t) \quad (8)$$

式中: $\Delta\tau_{ij}^k(t)$ 是 k 于此次循环中留在路径 (i, j) 上的信息量,其值由公式(9)表示。

4 信息数据节点判定准则

在大数据中,不仅仅需要识别出信息数据在目

标事件的作用^[4],以及信息数据节点之间通过该事件形成的关系^[5],更需要分辨出这些网络信息数据在该事件中关系的重要性^[6]。信息数据节点的重要性或关键性的判定可以通过节点关系的点度中心度、紧致中心度和节点活跃度进行综合判定。

1) 点度中心度^[7] (Degree Centrality)。信息数据节点的点度中心度指与其相连的信息节点数。即:

$$d_i = \sum_{j=1}^n e_{i,j} \quad (10)$$

$$e_{i,j} = \begin{cases} 1 & \text{节点 } v_i \text{ 和 } v_j \text{ 间有边相连} \\ 0 & \text{其他} \end{cases}$$

点度中心度反映了信息数据节点在网络中的活跃程度,刻画了信息数据节点对目标事件关系的影响力。点度中心度越高,表示与该节点连接的节点数越多,说明该信息数据节点越有能力支配其它节点。

2) 紧致中心度 (Closeness Centrality)。网络信息数据节点的紧致中心度是指它与网络中所有其它节点的最短路径长度和。

$$c_i = \sum_{j=1}^n p_{i,j} \quad (11)$$

式中: $p_{i,j}$ 为节点 v_i 和 v_j 间最短路径的长度。

紧密中心度反映了信息数据节点到其它信息数据节点间的平均路径长度。刻画了该节点居于网络中心位置的能力以及从该节点到达其它节点的速率。紧致中心度越小,表示该节点越处于网络的中间位置,通过该节点向其它节点发送信息速率越快。

3) 节点活跃度 (Activity)。信息数据节点的活跃度等于与其连接的所有边的信息素之和。即:

$$a_i = \sum_{j=1}^n \text{Flow}[i][j] \quad (12)$$

节点活跃度反应了与之直接关联的节点集中的重要程度,在信息数据网络中,选择累积活跃度达到总活跃度一定比例(比例因子为 r)的节点为重要目标节点集。其数学模型为:从包含 n 个节点 $V = \{v_1, v_2, \dots, v_n\}$ 的网络中,选出 m 个节点 $V' = \{v_1, v_2, \dots, v_m\}$, $V'_k \in V'$, 设节点 V'_i 的活跃度为 a_i ,使 V' 满足:

$$\sum_{i=1}^m a'_i \geq r \times \sum_{i=1}^n a_i \quad (13)$$

式中:比例因子 r 可以根据具体情况设定,一般取 $r \in [0.5, 1]$ 。

在目标事件中,根据信息数据的重要性或关键性,将信息数据分为关键目标信息、重要目标信息和一般目标信息 3 类,则根据点度中心度、紧致中心度和目标节点活跃度的定义,可以定义网络信息数据识别的准则。

判定准则 1: 在某个目标信息数据节点关系中,如果一个目标节点 v 属于重要目标节点集,并且该目标节点的点度中心度最大或紧致中心度最短,则该目标信息数据节点是关键目标信息。

判定准则 2: 在某个目标信息数据节点关系中,如果一个目标信息数据节点 v 属于重要目标节点集,但该目标节点的点度中心度不是最大,同时紧致中心度不是最短,则该目标信息数据节点是重要目标信息。

判定准则 3: 在某个目标信息数据节点关系中,如果一个目标信息数据节点 v 不属于重要目标节点集,则该目标信息数据节点为一般目标。

5 仿真实验

本仿真试验中的目标节点包括关键目标节点与一般目标节点:关键目标节点是包含 1 个及以上重要目标信息数据的节点,一般目标节点是只包含一般目标信息数据的节点。若干重要目标信息数据及若干一般目标信息数据构成了若干关键目标节点与若干一般目标节点,进而组成了若干网络节点。

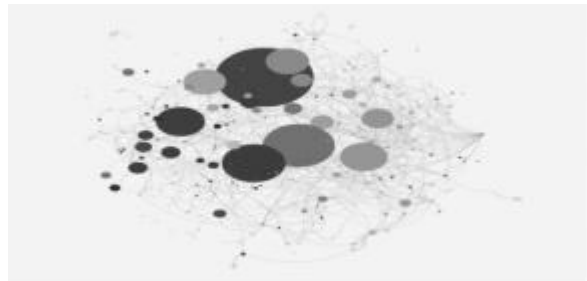


图 3 信息数据结构划分图

Fig. 3 Structure partition diagram of information data
将 280 000 个由信息数据以 FHACC^[8] 进行数据结构划分,如图 3 所示,所有的点都表示在信息网络中的节点,点的大小表示了该网络节点包含的

信息数据量,不同的颜色表示不同的网络位置,同颜色不同大小的点表示不同的网络节点所在位置相同。

如图 4 所示,同一网络时间条件下,三角形表示数据中已知的目标事件一般目标节点,圆点表示数据中已知的目标事件关键目标节点,点的大小表示该节点所含信息量的多少。

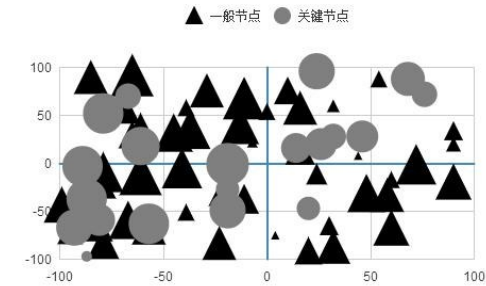


图 4 目标事件一般目标节点与关键目标节点初始分布情况图

Fig. 4 Initial distribution diagram of general target nodes and key target nodes in target event

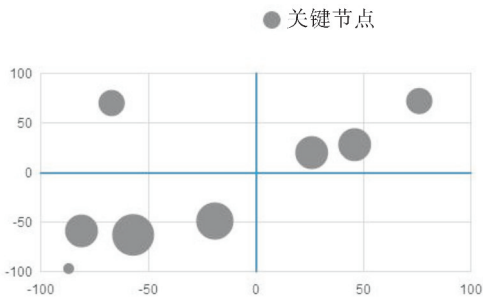


图 5 搜索出的关键目标节点分布图

Fig. 5 Distribution diagram of detected key target nodes



图 6 HBACA 算法提取出的关系图

Fig. 6 Relation diagram extracted through HBACA algorithm

对 280 000 个信息数据进行聚类后,采用 HBACA 对网络信息数据关系进行识别,其计算结果可大致分辨出对每个信息数据的重要性,但无法找出其间的关联,以及其在目标事件中的作用。

HBACA 搜索出的关键目标节点分布情况,如图 5 所示。HBACA 重要目标节点输出数据以 FHACC 进行数据结构划分结果,如图 6 所示。

经过对 HBACA 算法进行优化后,本文设计的算法,根据计算输出结果,能够清楚地看到相关目标的关系非常明显如图 7,很容易获得所处事件中重要目标之间的关系如图 8 所示。

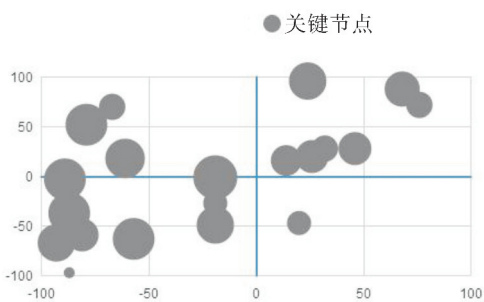


图 7 搜索出的关键目标节点分布图

Fig. 7 Distribution diagram of detected key target nodes



图 8 本文的方法提取出的关系图

Fig. 8 Relation diagram extracted by method proposed in this paper

图 8 所示,不同颜色代表不同的目标分类。

由图 3-8 对比可知,基于关系的目标信息数据识别方法在同一目标事件中识别出的关键目标节点,与已知的关键目标节点数量一致;基于关系的信息数据识别方法能够识别出同一目标事件中的节点,而且识别出的网络节点数量比 HBACA 算法识别出的关键目标节点要多;本文基于关系的信息数据识别方法比 HBACA 算法的优势在于能够识别出同一目标事件信息数据之间的关系。

6 结束语

本文的算法,根据计算数据输出结合 FHACC 进行数据结构划分结果,能够明确地获得关键目标的信息数据坐标,并与其所处事件的重要程度能够

匹配,同时相关信息数据的关系非常明显,很容易提炼出所处目标事件中重要目标信息数据之间的关联。

参考文献

[1] 于歌,朴胜春.强海洋混响环境下的移动多目标检测方法[J].华中科技大学学报(自然科学版),2017,45(03):89-93.

[2] 马勋,周长胜,吕学强,等.基于 SAO 结构的非分类关系抽取研究[J].计算机工程与应用,2018(8): 25-27.

[3] 周袅,葛洪伟,苏树智.基于信息素的自适应连续域混合蚁群算法[J]. 计算机工程与应用,2017,53(6):

156-161.

[4] 李燕梅. 云计算背景下信息数据保护的关键技术研究[J]. 网络安全技术与应用, 2018(1):12-14.

[5] 赵逸智,张云峰. 大数据环境下文本信息挖掘系统设计[J]. 现代电子技术,2018,41(1):125-128.

[6] 朱荣,周彩兰,高瑞. 基于数据挖掘的客户关系管理系统研究[J]. 现代电子技术,2018,41(1):182-186.

[7] 钱珺,王朝坤,郭高扬. 基于社区的动态网络节点介数中心度更新算法[J]. 软件学报, 2018(3):56-57.

[8] SHIROKOV M E. On discontinuity of information characteristics of quantum systems and channels[J]. Problems of Information Transmission, 2016, 52(3): 17-18.

(上接第 54 页)

[3] 潘光,黄明明,宋保维,等.AUV 回收技术现状及发展趋势[J].鱼雷技术,2008,16(6):10-14.

[4] 盛振邦,刘应中.船舶原理(下)[M].上海:上海交通大学出版社,2004.

[5] WHEELER D, WILKINSON C. From calm to storm; the origins of the Beaufort wind scale[J]. The Mariner's Mirror, 2004, 90(2): 187-201.

[6] 许景波,边信黔,付明玉.随机海浪的数值仿真与频谱分析[J].计算机工程与应用,2010,46(36):226-229.

[7] 徐德伦.由 JONSWAP 谱和 PM 谱计算的风浪波高之间的关系[J].海洋湖沼通报,1987(1):1-4.

[8] 俞聿修.随机波浪及其工程应用[M].大连:大连理工大学出版社,2000.

[9] 张铁栋.潜水器设计原理[M].哈尔滨:哈尔滨工程大学出版社,2011.

[10] 罗敏莉,毛筱菲,王晓侠.强迫运动柱体附加质量与阻尼系数的 CFD 计算[J].水动力学研究与进展:A 辑, 2011,26(4):509-515.

[11] 胡志强,林扬,谷海涛.水下机器人粘性类水动力数值计算方法研究[J].机器人,2007,29(2):145-150.

[12] 刘献栋,邓志党,高峰.基于逆变换的路面不平度仿真研究[J].中国公路学报,2005,18(1):122-126.