УДК 621.396.96

Повышение вероятностных характеристик отождествления целевой информации накоплением принимаемых решений

Савченко Д. И.

Научно-производственная фирма «Меридиан», ул. Блохина, 9, Санкт-Петербург, 197198, Россия

e-mail: cfdX@list.ru

Аннотация

В статье предложены решающее правило и алгоритм, позволяющие повысить вероятностные характеристики отождествления целевых данных от двух источников информации за счет накопления и учета решений, принятых на нескольких вызовах задачи отождествления. Сложность предложенного алгоритма меньше, чем у известного аналога.

Ключевые слова: отождествление, источники информации, целевая обстановка, третичная обработка, информационно-управляющая система

Введение

Для получения актуальной информации об обстановке, используемой в информационно-управляющих системах (ИУС) для решения широкого круга задач, на современные носители устанавливается большое число средств, систем и комплексов обнаружения и сопровождения целей. На основе данных от всех источников информации (ИИ) путем решения задачи отождествления в ИУС формируется единая целевая обстановка, более полная и достоверная в сравнении с информацией отдельных ИИ [1, 2]. Под отождествлением понимается процесс принятия решений о принадлежности или непринадлежности данных о целях (формуляров целей), поступающих от отдельных ИИ, одним и тем же реальным объектам. В общем случае принятие таких решений носит вероятностный характер, поэтому на ряду с правильными могут возникать ложные отождествления и неотождествления формуляров целей (ФЦ) от ИИ. Это снижает эффективность решения задач на основе отождествленных данных об обстановке. Поэтому актуальной является разработка способов и алгоритмов, предназначенных для повышения вероятностных характеристик отождествления информации о целях.

Задача отождествления решается в ИУС регулярно, обычно, с некоторым постоянным периодом. Это дает возможность повышения вероятностных характеристик отождествления за счет накопления и учета решений о тождественности или нетождественности ФЦ, принятых на нескольких вызовах задачи отождествления.

Многие алгоритмы отождествления целевой информации от большого числа ИИ сводятся к многократным вызовам процедуры отождествления данных от пары ИИ [1, 2]. В этой связи, в настоящей работе также рассматривается алгоритм, позволяющий повысить вероятностные характеристики отождествления целевой информации от двух ИИ за счет учета решений, принятых на нескольких вызовах.

Алгоритмы, выполняемые для отождествления ФЦ от пары ИИ на каждом вызове задачи отождествления, далее обозначаемые АОД, рассмотрены, например, в [1, 2].

В работе [3] предложен алгоритм, аналогичный синтезируемому. Однако сложность данной процедуры составила $O\left((n_1+n_2)^3\right)$, где n_1, n_2 – максимальное количество ФЦ от первого и второго ИИ соответственно. С учетом того, что сложность АОД, описанных, например, в [1, 2], составляет $O\left(\max(n_1, n_2)^3\right)$, применение алгоритма из [3] существенно увеличивает время решения задачи отождествления в целом. Поэтому целесообразна модификация указанного алгоритма с целью понижения его сложности.

Постановка задачи

Рассмотрим результаты выполнения АОД на последних t вызовах задачи отождествления. Максимальное количество ФЦ от первого ИИ обозначим n_1, от второго – n_2. Для облегчения формальной записи введем дополнительный фиктивный ФЦ от второго ИИ с номером $z=-1$ и будем считать, что если i-ый ФЦ от первого ИИ после некоторого вызова АОД признан неотождествленным, то он отождествлен (на этом вызове) с z-ым ФЦ от второго ИИ. Тогда для каждого k-ого ФЦ от первого ИИ имеется последовательность из t номеров ФЦ от второго ИИ, с которыми он (k-ый ФЦ) был отождествлен на последних t вызовах АОД. Все такие последовательности образуют матрицу $W=\left\|w_{i,k}\right\|_{\substack{i=\overline{1,t}\\ k=\overline{1,n_1}}}$, где $w_{i,k}\in\{1,..,n_2\}\cup\{z\}$ – номер ФЦ от второго ИИ, с которым был отождествлен k-ый ФЦ от первого ИИ i вызовов АОД назад. Для каждого столбца этой матрицы выполняется условие:

$$\forall i = \overline{1,t},\ \forall k = \overline{1,n_1},\ \forall m = \overline{1,n_1},\ m \neq k,\ w_{i,k} \neq z : w_{i,k} \neq w_{i,m}. \qquad (1)$$

То есть на одном вызове АОД один ФЦ (реальный) от второго ИИ не может быть отождествлен более чем с одним ФЦ от первого ИИ. Фиктивный ФЦ от второго ИИ может быть отождествлен с произвольным числом ФЦ от первого ИИ, так как произвольное число ФЦ от первого ИИ могут быть признаны неотождествленными.

Пусть от первого ИИ имеется всего один ФЦ $(n_1 = 1)$. В этом случае выбор ФЦ от второго ИИ для отождествления синтезируемым алгоритмом (то есть по результатам t вызовов АОД) с единственным ФЦ от первого ИИ целесообразно осуществлять с использованием мажоритарного решающего правила [4]. В соответствие с этим правилом, из множества ФЦ от второго ИИ, с которыми отождествлялся единственный ФЦ от первого ИИ на последних t вызовах АОД, должен выбираться тот, с которым отождествление производилось наибольшее число раз. То есть из единственного столбца матрицы W должен быть выбран номер ФЦ от второго ИИ, встречающийся чаще других:

$$g' = \arg \max_{g \in \{1,..,n_2\} \cup \{z\}} \sum_{i=1}^{t} \beta\left(w_{i,1} = g\right), \qquad (2)$$

где

β – функция, возвращающая единицу при истинном значении своего аргумента (то есть при $w_{i,1} = g$) и ноль в противном случае.

При наличии нескольких номеров ФЦ от второго ИИ, встречающихся в столбце $\left\| w_{1,1}, w_{2,1}, \ldots, w_{t,1} \right\|^T$ одинаковое наибольшее число раз, выбор среди них, в силу равновероятности соответствующих событий в рамках рассматриваемой модели, произволен.

В общем случае от первого ИИ имеется до n_1 ФЦ и результатом работы рассматриваемого алгоритма должен быть не один номер g' ФЦ от второго ИИ, а упорядоченная последовательность $G_l = \{g_{1,l}, g_{2,l}, \ldots, g_{n_1,l}\}$, где $g_{i,l} \in \{1,..,n_2\} \cup \{z\}$ – номер ФЦ от второго ИИ, с которым отождествляется i-ый ФЦ от первого ИИ на основе решений АОД на последних t вызовах. Так как ФЦ (не считая фиктивного) от второго ИИ не может быть отождествлен более чем с одним ФЦ от первого ИИ, для любой допустимой последовательности G_l можно записать (по аналогии с (1)):

$$\forall i = \overline{1, n_1},\ \forall j = \overline{1, n_1},\ i \neq j,\ g_{i,l} \neq z : g_{i,l} \neq g_{j,l}\,. \qquad (3)$$

Другое ограничение на G_l формально может быть записано в виде:

$$\forall i = \overline{1, n_1},\ g_{i,l} \neq z : \sum_{j=1}^{t} \beta(w_{j,i} = g_{i,l}) > 0\,. \qquad (4)$$

Условие (4) означает, что в соответствии с допустимой последовательностью G_l два ФЦ от разных ИИ не могут быть отождествлены синтезируемым алгоритмом, если они ни разу не отождествились на последних t вызовах АОД. Если не существует ФЦ от второго ИИ, с которым i-ый ФЦ от первого ИИ мог бы быть отождествлен, то должно приниматься решение о нетождественности последнего (то есть об отождествлении с фиктивным z-ым ФЦ от второго ИИ).

Введем понятие веса последовательности G_l, определяемого по следующей формуле:

$$S(G_l) = \sum_{i=1}^{n_1} \sum_{j=1}^{t} \beta(w_{j,i} = g_{i,l})\,. \qquad (5)$$

Решающее правило для данного общего случая может быть получено обобщением правила (2): в качестве итоговой синтезируемым алгоритмом должна выбираться допустимая последовательность G_l, обладающая максимальным весом:

$$G' = \arg \max_{G_l} S(G_l) = \arg \max_{G_l} \sum_{i=1}^{n_1} \sum_{j=1}^{t} \beta(w_{j,i} = g_{i,l})\,. \qquad (6)$$

На практике значение t должно определяться для каждой конкретной ИУС исходя из требований к максимальной задержке между моментами взятия цели на сопровождение и принятия решения о тождественности, которая при использовании решающего правила (6) в самых благоприятных условиях составляет $]0.5 \cdot t[\cdot T$, где T – период решения задачи отождествления.

Синтез алгоритма

Искомая последовательность G' обладает важным для построения алгоритма свойством:

$$\forall i = \overline{1, n_1} : \sum_{j=1}^{t} \beta(w_{j,i} = g_i') \geq \sum_{j=1}^{t} \beta(w_{j,i} = z), \tag{7}$$

где

g_i' – i-ый элемент последовательности G'.

Условие (7) означает, что если два ФЦ от разных ИИ на последних t вызовах АОД отождествились меньшее число раз, чем ФЦ от первого ИИ из этой пары признавался неотождествленным, то синтезируемым алгоритмом не может быть принято решение об отождествлении этой пары ФЦ. Действительно, если бы условие (7) не выполнялось, то заменой в последовательности G' значения g_i', не отвечающее (7), на z, могла бы быть получена последовательность, удовлетворяющую свойствам (3) и (4), но с большим (чем G') весом, что противоречит (6).

Возможно наличие нескольких оптимальных в смысле (6) и эквивалентных с точки зрения синтезируемого алгоритма последовательностей G'. В частности, в G' могут существовать элементы $g_i' \neq z$, для которых в (7) имеет место равенство. Тогда существует (как минимум одна) полностью аналогичная допустимая последовательность того же веса, в которой $g_i' = z$. Условимся в дальнейшем из таких оптимальных в смысле (6) последовательностей всегда выбирать вторую. Тогда условие (7) примет вид:

$$\forall i = \overline{1, n_1},\ g_i' \neq z : \sum_{j=1}^{t} \beta(w_{j,i} = g_i') > \sum_{j=1}^{t} \beta(w_{j,i} = z), \tag{8}$$

В контексте решаемой задачи это означает, что если ФЦ от первого ИИ на последних t вызовах АОД отождествился с некоторым реальным ФЦ от второго ИИ наибольшее число раз и столько же раз признавался неотождествленным, то синтезируемый алгоритм примет решение о нетождественности данного ФЦ от первого ИИ.

Так как $\sum_{j=1}^{t} \beta(w_{j,i} = z) \geq 0$, то условие (8) является более сильным, чем условие (4).

С учетом условия (8), вес оптимальной последовательности может быть записан в форме:

$$S(G') = \sum_{i=1}^{n_1} \sum_{j=1}^{t} \beta(w_{j,i} = g'_{i,l}) = \sum_{i=1}^{n_1} \max\left(\sum_{j=1}^{t} \beta(w_{j,i} = g'_{i,l}), \sum_{j=1}^{t} \beta(w_{j,i} = z) \right). \tag{9}$$

Пусть $n_1 \le n_2$. Введем в рассмотрение двудольный граф, первая доля которого образована n_1 вершинами, отвечающими ФЦ от первого ИИ, вторая доля – n_2 вершинами, отвечающими ФЦ от второго ИИ, а веса ребер между каждой парой вершин (i, j) разных долей определяются следующим образом:

$$r_{i,j} = \begin{cases} \sum_{m=1}^{t} \beta(w_{m,i} = z), \sum_{m=1}^{t} \beta(w_{m,i} = j) \le \sum_{m=1}^{t} \beta(w_{m,i} = z) \\ \sum_{m=1}^{t} \beta(w_{m,i} = j), \sum_{m=1}^{t} \beta(w_{m,i} = j) > \sum_{m=1}^{t} \beta(w_{m,i} = z) \end{cases} = \begin{cases} q_{i,z}, \ q_{i,j} \le q_{i,z} \\ q_{i,j}, \ q_{i,j} > q_{i,z} \end{cases} = \max(q_{i,z}, q_{i,j}), \tag{10}$$

где

$q_{i,j} = \sum_{m=1}^{t} \beta(w_{m,i} = j)$ – количество повторений номера j в i-ом столбце матрицы W.

Рассмотрим совершенное паросочетание M в этом двудольном графе. Так как количество вершин первой доли не больше количества вершин второй доли ($n_1 \le n_2$), то M, по определению совершенного паросочетания, включает n_1 ребер и покрывает все вершины первой доли. Если упорядочить ребра из M по возрастанию номеров вершин первой доли, то формула для его веса может быть записана в виде:

$$S(M) = \sum_{i=1}^{n_1} r_{i,m_i} = \sum_{i=1}^{n_1} \max(q_{i,z}, q_{i,m_i}), \tag{11}$$

где

m_i ($i = \overline{1, n_1}$) – вершина второй доли, инцидентная i-му ребру из паросочетания M.

Таким образом, вес произвольного совершенного паросочетания M дает значение веса последовательности G_M, элементы которой определяются как:

$$g_{i,M} = \begin{cases} m_i, & q_{i,m_i} > q_{i,z} \\ z, & q_{i,m_i} \le q_{i,z} \end{cases} = \begin{cases} m_i, & q_{i,m_i} > q_{i,z} \\ z, & q_{i,m_i} = q_{i,z} \end{cases}. \tag{12}$$

Для проверки того, что G_M является допустимой последовательностью необходимо убедиться в выполнении для нее условий (3) и (8). Выполнение условия (3) вытекает из самого определения паросочетания: никакие два входящие в него ребра не могут быть инциденты одной вершине, следовательно:

$$\forall i = \overline{1, n_1}, \forall j = \overline{1, n_1}, i \neq j : m_i \neq m_j \Rightarrow \forall i = \overline{1, n_1}, \forall j = \overline{1, n_1}, i \neq j, g_{i,M} \neq z : g_{i,M} \neq g_{j,M}. \qquad (13)$$

Выполнение условия (8) вытекает из способа (12) формирования последовательности на основе совершенного паросочетания M.

Таким образом, любое совершенное паросочетание M в рассматриваемом двудольном графе отвечает некоторой допустимой последовательности G_M, а вес M совпадает с весом этой последовательности. Кроме того, оптимальная (в смысле (6)) последовательность G', удовлетворяющая условиям (3) и (8), имеет в рассматриваемом двудольном графе соответствующее ей совершенное паросочетание, которое может быть построено, например, следующим образом:

$$m_i = \begin{cases} g_{i,l}, & g_{i,l} \neq z \\ k_i, & g_{i,l} = z \end{cases}, \qquad (14)$$

где

$k_i \in \{x \mid x \in \{1,..,n_2\}, x \notin \{g_{j,l} \mid j = \overline{1, n_2}, g_{j,l} \neq z\}\}$; $i \neq j \rightarrow k_i \neq k_j$.

Наличие достаточного числа элементов k_i в (14) следует из того, что $n_1 \leq n_2$. Причем веса ребер (i, k_i) в графе будут равны $q_{i,z}$ (легко показать, что противное говорило бы о неоптимальности последовательности), а значит по формуле (12) из полученного совершенного паросочетания может быть сформирована исходная последовательность G'.

Из сказанного следует, что для случая $n_1 \leq n_2$ искомая последовательность G' может быть найдена путем поиска совершенного паросочетания максимального веса в рассмотренном двудольном графе и применения формулы (12).

В случае $n_1 > n_2$ совершенное паросочетание в рассмотренном двудольном графе не будет покрывать все вершины первой доли и полученные выводы не будут верны. Для исправления этой ситуации добавим во вторую долю $(n_2 - n_1)$ фиктивных вершин (в итоге количество вершин в обеих долях будет равно n_1). Веса ребер между каждой добавленной вершиной второй доли и каждой вершиной первой доли определим формулой:

$$r_{i,j} = q_{i,z} = \sum_{m=1}^{t} \beta\left(w_{m,i} = z\right), \quad j > n_2 . \tag{15}$$

Так как $q_{i,z} \geq 0$, а $\forall j > n_2 : q_{i,j} = 0$, то, с учетом (10) и (15), вес произвольного ребра рассматриваемого двудольного графа будет определяться формулой:

$$r_{i,j} = \begin{cases} \max(q_{i,z}, q_{i,j}), & j \leq n_2 \\ q_{i,z}, & j > n_2 \end{cases} = \begin{cases} \max(q_{i,z}, q_{i,j}), & j \leq n_2 \\ \max(q_{i,z}, q_{i,j}), & j > n_2 \end{cases} = \max(q_{i,z}, q_{i,j}) . \tag{16}$$

Это позволяет обобщить приведенные выше выводы и на случай $n_1 > n_2$.

В контексте решаемой задачи, добавление вершин во вторую долю отвечает введению дополнительных ФЦ от второго ИИ. Причем ФЦ от первого ИИ ни разу не отождествлялись с этими дополнительными ФЦ от второго ИИ. Поэтому номера дополнительных ФЦ не могут иметь место в выходной последовательности в силу (8). Таким образом, их введение никак не сказывается на получаемом результате, но позволяет обобщить алгоритм, корректный для $n_1 \leq n_2$, и на случай $n_1 > n_2$.

В итоге, синтезируемый алгоритм принимает следующий вид:

1) на основе матрицы W построить двудольный граф с n_1 вершинами в первой доле и $\max(n_1, n_2)$ вершинами во второй. Вес ребра между вершинами i и j соответственно первой и второй доли этого графа определить по формуле:

$$r_{i,j} = \max(q_{i,z}, q_{i,j}) . \tag{17}$$

2) найти совершенное паросочетание максимального веса M в двудольном графе. Для этого может быть использован Венгерский метод решения задачи о назначениях [5];

3) на основе найденного паросочетания M (упорядоченного в порядке возрастания номеров вершин первой доли) построить последовательность G, i-ый элемент которой определяется по формуле:

$$g_i = \begin{cases} m_i, & q_{i,m_i} > q_{i,z} \\ z, & q_{i,m_i} = q_{i,z} \end{cases} = \begin{cases} m_i, & (q_{i,m_i} > q_{i,z})\,\&\,(1 \le m_i \le n_2) \\ z, & (q_{i,m_i} = q_{i,z}) \vee (1 > m_i) \vee (m_i > n_2) \end{cases}. \quad (18)$$

4) признать отождествленными (по результатам решений АОД на последних t вызовах) пары ФЦ от первого и второго ИИ с номерами i и j, для которых

$$g_i = j,\ j \ne z. \quad (19)$$

Сложность первого шага составляет $O(n_1 \cdot (n_2 + t))$, второго – $O(\max(n_1, n_2)^3)$ (в соответствии с [5]), третьего и четвертого шагов – $O(n_1)$. Так как на практике $t < n_2$, сложность всего предложенного алгоритма составляет $O(\max(n_1, n_2)^3)$, что меньше, чем в ранее известном подходе [3] ($O((n_1 + n_2)^3)$), и соответствует сложности самого АОД [1, 2].

Результаты экспериментального исследования

Оценка адекватности предложенного алгоритма производилась посредством математического моделирования. По аналогии с [1], для отождествления использовалась информация о пеленгах и дистанциях до целей. Объекты располагались равномерно в пространстве признаков, причем для любых двух близлежащих объектов расстояние по одному признаку устанавливалось равным нулю, а по второму – $d \cdot \sigma_v$, где σ_v – суммарная среднеквадратическая погрешность определения соответствующего признака ИИ, d – коэффициент. При моделировании задавались различные значения других параметров (количество целей, вероятности взятия целей на сопровождение ИИ и др.). В качестве АОД использовался алгоритм [1]. Значение t устанавливалось равным 5.

В результате моделирования было установлено, что применение предложенного алгоритма действительно повышает вероятностные характеристики отождествления целевой информации в большинстве ситуаций. Так, например, вероятности принятия правильных решений выше 0,75..0,80 достигаются уже при расстояниях между целями хотя бы по одному признаку, в 1,4..1,6 раза превышающих суммарную среднеквадратическую погрешность определения соответствующего признака ИИ, в то время как сам АОД (без использования синтезированного алгоритма) обеспечивает указанные значения вероятностей только начиная с 1,6..1,9 раз [1].

Заключение

В статье предложены решающее правило и основанный на нем алгоритм, обеспечивающий повышение вероятностных характеристик отождествления целевой информации за счет накопления и учета решений о тождественности и нетождественности ФЦ от двух ИИ за несколько вызовов задачи отождествления. Сложность синтезированного алгоритма составила $O(\max(n_1, n_2)^3)$, что меньше, чем у известного аналога, и соответствует сложности самих АОД, используемых для отождествления на каждом вызове. Таким образом, применение рассмотренной процедуры позволяет повысить вероятности правильных отождествлений и неотождествлений без увеличения сложности всего алгоритма отождествления, что обосновывает целесообразность ее практического использования.

Библиографический список

1. Савченко Д.И. Оптимальные решающие правила и алгоритм отождествления целевой информации // Электронный журнал «Труды МАИ». – 2012. – № 58, URL: http://www.mai.ru/science/trudy/published.php?ID=33244 (дата обращения: 28.09.2012).

2. Савченко Д.И. Отождествление целевых данных в информационно-управляющих системах надводных кораблей и судов // Современные проблемы науки и образования. – 2012. – № 4 – URL: www.science-education.ru/104-6773 (дата обращения: 27.07.2012).

3. Савченко Д.И. Повышение качества отождествления целевой информации применением накапливающего алгоритма // Сборник трудов XVII международной научно-практической конференции студентов, аспирантов и молодых учёных "Современные техника и технологии". – Томск: ТПУ, 2011. – Т. 2. – С. 413-415.

4. Намичейшвили О., Элизбарашвили А. Математическая теория резервирования на основе модели формального нейрона // Georgian Electronic Scientific Journal: Computer Science and Telecommunications – 2009. – №4 (21).-136-189.

5. Пападимитриу Х., Стайглиц К. Комбинаторная оптимизация. Алгоритмы и сложность. – М: Мир, 1984.- 510 с.