УДК 621.396.96

Оптимальные решающие правила и алгоритм отождествления целевой информации

Д.И. Савченко

Аннотация

В статье рассмотрен подход к отождествлению данных об обстановке от двух источников информации. Предложены оптимальные универсальное и частное решающие правила и соответствующий полиномиальный алгоритм отождествления. При синтезе решающих правил снят ряд допущений известных правил.

Ключевые слова

отождествление; источники информации; целевая обстановка; третичная обработка; системы освещения обстановки

Информация о целевой обстановке в районе нахождения носителя является одной из основных составляющих исходных данных для широкого спектра задач в автоматизированных информационно-управляющих системах различного назначения (АИУС). По этой причине полнота и достоверность освещения обстановки ключевым образом влияют на эффективность решения этих задач и работу АИУС в целом. Для получения данных о воздушных и надводных объектах современные носители оснащаются большим числом источников информации (ИИ). Эти обстоятельства обуславливают необходимость решения задачи отождествления информации от всех ИИ, то есть принятия решений о принадлежности и непринадлежности данных о целях (формуляров целей), поступивших от разных ИИ, одним и тем же реальным объектам. Формируемая задачей отождествления единая картина целевой обстановки является более полной и достоверной по сравнению с информацией от каждого из ИИ в отдельности.

Состояние рассматриваемого вопроса. Постановка задачи

При разработке алгоритмов отождествления должен учитываться ряд факторов, вытекающих из существующих проблем предметной области:

– рост количества целей и числа ИИ, данные от которых поступают на обработку в АИУС. С учетом необходимости решения задачи отождествления в реальном времени, это накладывает существенные ограничения на сложность используемых алгоритмов, которые можно считать практически применимыми, если время их выполнения полиномиально зависит от количества сопровождаемых ИИ целей [1]. Учитывая типовые размеры исходных данных и возможности современных вычислителей, степень полинома не должна превышать 3-4;

– соизмеримость расстояний между целями и среднеквадратических ошибок (СКО) выработки признаков (координат, параметров движения и т.п.) целей современными ИИ, что приводит к пересечению областей вероятного нахождения многих целей.

Построение алгоритма отождествления целевой информации от большого числа ИИ обычно осуществляется путем многократных вызовов процедур отождествления данных от пар ИИ. Поэтому основное внимание во многих работах [2, 3, 4, 5, 6, 7], посвященных этой задаче, в том числе и в настоящей, уделено вопросам отождествления данных от двух ИИ.

При построении алгоритмов отождествления информации от пары ИИ широкое применение находит двухэтапная схема [3, 4, 5, 6], позволяющая без существенного уменьшения вероятностей принятия правильных решений (ВПР) сократить время вычислений.

На первом этапе сравнением разностей признаков для отождествления из всех пар формуляров целей (ФЦ) от двух ИИ с порогами, вычисляемыми на основе СКО по этим признакам (обычно по правилу «трех сигм»), принимаются решения о предварительных отождествлениях ФЦ. Под ФЦ подразумевается структура, включающая L-вектор X выработанных ИИ признаков для отождествления цели и корреляционную матрицу Ψ ошибок его определения.

На втором этапе на основе результатов первого этапа формируются непересекающиеся группы предварительно отождествленных ФЦ (ГФЦ) и производится отождествление внутри каждой из них.

Алгоритм выполнения первого этапа тривиален. Выделение ГФЦ сводится к выделению компонент связности двудольного графа, множества вершин долей которого отвечают множествам ФЦ от ИИ, а наличие или отсутствие ребра между каждой парой вершин определяется выполнением условия предварительного отождествления для соответствующих ФЦ [4]. Таким образом, основной интерес представляет алгоритм отождествления данных внутри ГФЦ, который и определяет временные и вероятностные характеристики всего алгоритма отождествления данных от двух ИИ.

Отождествление внутри ГФЦ – это процесс выбора одной из непротиворечивых совокупностей гипотез о принадлежности или непринадлежности всех пар входящих в ГФЦ формуляров целей от разных ИИ одним и тем же реальным объектам. Далее такие совокупности гипотез условимся называть вариантами отождествления (ВО) внутри ГФЦ.

Исходными данными $U=\{U_1,U_2\}$ для алгоритма отождествления внутри ГФЦ являются наборы включенных в нее ФЦ от двух ИИ: $U_1=\{u_{1,1},u_{1,2}...u_{1,N_1}\}$ и $U_2=\{u_{2,1},u_{2,2}...u_{2,N_2}\}$, где N_1, N_2 – количество ФЦ от первого и второго ИИ в ГФЦ, $u_{i,j}$ – j-ый ФЦ в ГФЦ от i-ого ИИ.

Количество ВО в ГФЦ определяется следующим образом [5]:

$$K=\sum_{n=0}^{\min(N_1,N_2)}\frac{1}{n!}\cdot\frac{N_1!}{(N_1-n)!}\cdot\frac{N_2!}{(N_2-n)!}. \qquad (1)$$

Любой допустимый ВО формально может быть описан множеством из $k_g\in\{0,..,\min(N_1,N_2)\}$ пар номеров ФЦ, отождествляемых в соответствии с этим ВО:

$$A_g\to\{y_{g,1},y_{g,2}...y_{g,k_g}\}=\{y_{g,m}\}_{m=1..k_g}, \qquad (2)$$

где

A_g – g-ый ВО внутри ГФЦ ($g=\overline{1,K}$);

$y_{g,m}=\{y^g_{1,m},y^g_{2,m}\}$ – пара номеров ФЦ от первого и второго ИИ, отождествляемых в соответствии с ВО A_g;

$y^g_{1,m}\in\{1,..,N_1\}$ – номер ФЦ от первого источника ($\forall i\neq m: y^g_{1,m}\neq y^g_{1,i}$);

$y^g_{2,m}\in\{1,..,N_2\}$ – номер ФЦ от второго источника ($\forall i\neq m: y^g_{2,m}\neq y^g_{2,i}$).

Известен [3, 4] подход к построению алгоритма отождествления внутри ГФЦ, основанный на предположении, что число целей в области пространства, занимаемой ГФЦ, равно числу ФЦ от ИИ, выдавшего большее их количество в ГФЦ. Тогда выбор осуществляется только из ВО, для которых $k_g=\min(N_1,N_2)$, а оптимальное по критерию максимума правдоподобия решающее правило (РП) имеет вид:

$$\Omega(U) = \arg\max_{A_g} \sum_{m=1}^{\min(N_1,N_2)} \Delta X_{y_{g,m}}^{T} \cdot \Psi^{\Sigma -1} \cdot \Delta X_{y_{g,m}}, \qquad (3)$$

где

$\Omega(U)$ – РП, по входным данным U возвращающее ВО;

$\Delta X_{y_{g,m}} = X_{1,y_{1,m}^{g}} - X_{2,y_{2,m}^{g}}$ – вектор разностей соответствующих признаков для отождествления из ФЦ с номерами $y_{1,m}^{g}$ и $y_{2,m}^{g}$ от первого и второго ИИ;

$\Psi^{\Sigma} = \Psi_1 + \Psi_2$ – суммарная корреляционная матрица ошибок измерения признаков для отождествления источниками (предполагается, что $\Psi_{i,1} = ... = \Psi_{i,N_i} = \Psi_i$, $i = 1,2$).

Предположения, сделанные в данном подходе, на практике обычно не выполняются, что приводит к существенному уменьшению ВПР.

Решающее правило и алгоритм, построенные без указанных допущений, предложены, например, в [5]. При их построении предполагалось, что плотность вероятности (p_M) нахождения любой цели в каждой точке области пространства признаков, где локализована ГФЦ, постоянна и не зависит положения других целей. Тогда апостериорная вероятность варианта отождествления A_g определяется выражением:

$$P(A_g \mid U) = \frac{C \cdot (N_1 + N_2 - k_g)! \cdot (q_2 \cdot q_1)^{N_1+N_2-k_g}}{(1 - q_2 \cdot q_1)^{N_1+N_2-k_g+1} \cdot p_M^{-\min(N_1,N_2)}} \cdot \prod_{m=1}^{k_g} \frac{\exp\left(-\frac{1}{2} \cdot \Delta X_{y_{g,m}}^{T} \cdot \Psi_{y_{g,m}}^{\Sigma\ -1} \cdot \Delta X_{y_{g,m}} \right)}{p_M \cdot (2 \cdot \pi)^{\frac{L}{2}} \cdot \sqrt{\left| \Psi_{y_{g,m}}^{\Sigma} \right|}}, \qquad (4)$$

где

q_1, q_2 – вероятности пропуска цели в рассматриваемой области пространства первым и вторым ИИ соответственно.

$\Psi_{y_{g,m}}^{\Sigma} = (\Psi_{1,y_{1,m}^{g}} + \Psi_{2,y_{2,m}^{g}})$ – суммарная корреляционная матрица ошибок из ФЦ с номерами $y_{1,m}^{g}$ и $y_{2,m}^{g}$ от первого и второго ИИ;

C – постоянный для всех ВО коэффициент.

Предложенный в [5] алгоритм выбирает ВО с максимальным значением апостериорной вероятности (4).

Предположение о независимости плотности вероятности нахождения цели в каждой точке пространства признаков от положения других целей приводит к игнорированию априорной информации о возможных пространственных отношениях между целями, что снижает эффективность практического применения этого подхода к отождествлению внутри ГФЦ.

Таким образом, целесообразен синтез новых РП и соответствующего алгоритма, лишенных указанных недостатков известных подходов.

Синтез универсального решающего правила

На основе теоремы Байеса оптимальное по критерию максимума апостериорной вероятности РП в общем виде может быть записано как:

$$\Omega(U) = \arg \max_{A_g, g=\overline{1,K}} \left[\frac{P(A_g) \cdot p(U \mid A_g)}{\sum_{g=1}^{K} P(A_g) \cdot p(U \mid A_g)} \right], \tag{5}$$

где

$P(A_g)$ – априорная вероятность ВО A_g;

$p(U \mid A_g)$ – плотность вероятности данных U при условии правильности ВО A_g.

Априорная вероятность ВО A_g определяется как вероятность взятия на сопровождение источниками k_g одних и тех же целей в области пространства, отвечающей ГФЦ, при условии сопровождения ими соответственно N_1 и N_2 целей в этой области, отнесенная к общему количеству ВО, в соответствии с которыми отождествляется k_g пар ФЦ. Выражение для априорной вероятности ВО имеет вид [5]:

$$P(A_g) \approx C_1 \cdot \frac{(N_1 + N_2 - k_g)! \cdot (q_2 \cdot q_1)^{N_1+N_2-k_g}}{(1-q_2 \cdot q_1)^{N_1+N_2-k_g+1}}, \tag{6}$$

где

C_1 – постоянный для всех ВО коэффициент.

С учетом независимости измерений по целям разными ИИ, и измерений одного ИИ по разным целям, условная плотность вероятности $p(U \mid A_g)$ в общем виде определяется следующим образом:

$$p(U \mid A_g) = \prod_{i=1}^{N_1} \prod_{j=1}^{N_2} p(u_{1,i}, u_{2,j} \mid h_{i,j}^g) = C_2 \cdot \prod_{m=1}^{k_g} \mu(u_{1,y_{1,m}^g}, u_{2,y_{2,m}^g}), \quad (7)$$

где

$h_{i,j}^g \in \{h_0, h_1\}$ – гипотеза, принимаемая в отношении i-ого ФЦ от первого ИИ и j-ого ФЦ от второго ИИ в ГФЦ в соответствии с ВО A_g: h_0 – принадлежат разным реальным целям; h_1 – принадлежат одной реальной цели;

$p(u_{1,i}, u_{2,j} \mid h_{i,j}^g)$ – условная плотность вероятности измерений $u_{1,i}$ и $u_{2,j}$ при гипотезе $h_{i,j}^g$;

$C_2 = \prod_{i=1}^{N_1} \prod_{j=1}^{N_2} p(u_{1,i}, u_{2,j} \mid h_0)$ – постоянный для любого ВО коэффициент;

$\mu(u_{1,y_{1,m}^g}, u_{2,y_{2,m}^g}) = \dfrac{p(u_{1,y_{1,m}^g}, u_{2,y_{2,m}^g} \mid h_1)}{p(u_{1,y_{1,m}^g}, u_{2,y_{2,m}^g} \mid h_0)}$ – отношение условных плотностей вероятностей измерений $u_{1,y_{1,m}^g}$ и $u_{2,y_{2,m}^g}$ при гипотезах об их принадлежности одной и разным целям.

Условная плотность вероятности получения данных $u_{1,i}$ и $u_{2,j}$ при гипотезе о принадлежности этих ФЦ разным реальным целям исходя из формулы полной вероятности может быть записана как:

$$p(u_{1,i}, u_{2,j} \mid h_0) = \int_{\Phi} p(u_{1,i}, u_{2,j} \mid h_0, M_1) \cdot p(M_1 \mid \widehat{U}) \cdot dM_1, \quad (8)$$

где

M_1 – вектор, задающий точку в пространстве признаков;

$\widehat{U}$ – вся совокупность данных по целям от источников, в общем случае включающая не только данные U по рассматриваемой ГФЦ;

$p(M_1 \mid \widehat{U})$ – плотность вероятности нахождения цели в каждой точке пространства признаков для отождествления;

$p(u_{1,i},u_{2,j}\,|\,h_0,M_1)$ – плотность вероятности получения данных $u_{1,i}$ и $u_{2,j}$ при условиях принадлежности этих ФЦ разным целям и нахождения первой цели в точке M_1 пространства признаков;

Ф – множество всех возможных положений цели в пространстве признаков (для рассматриваемой ГФЦ).

Аналогично, плотность $p(u_{1,i},u_{2,j}\,|\,h_0,M_1)$ определяется как:

$$p(u_{1,i},u_{2,j}\,|\,h_0,M_1)=\int_{\Phi} p(u_{1,i},u_{2,j}\,|\,M_1,M_2)\cdot p(M_2\,|\,M_1,\widehat{U})\cdot dM_2, \tag{9}$$

где

$p(u_{1,i},u_{2,j}\,|\,M_1,M_2)$ – плотность вероятности получения данных $u_{1,i}$ и $u_{2,j}$ при нахождении соответствующих целей в точках M_1 и M_2;

$p(M_2\,|\,M_1,\widehat{U})$ – плотность вероятности нахождения цели в точке M_2 пространства признаков при условии нахождения другой цели в точке M_1.

С учетом нормальности закона распределения ошибок ИИ и независимости измерений, $p(u_{1,i},u_{2,j}\,|\,M_1,M_2)$ имеет вид:

$$p(u_{1,i},u_{2,j}\,|\,M_1,M_2)=\gamma(X_{1,i},M_1,\Psi_{1,i})\cdot\gamma(X_{2,j},M_2,\Psi_{2,j}), \tag{10}$$

где

$$\gamma(X,M,\Psi)=\frac{\exp\left(-\frac{1}{2}\left(X-M\right)^T\cdot\Psi^{-1}\cdot\left(X-M\right)\right)}{(2\cdot\pi)^{\frac{L}{2}}\cdot\sqrt{|\Psi|}}.$$

Условная плотность вероятности измерений $u_{1,i}$ и $u_{2,j}$ при гипотезе об их принадлежности одной реальной цели может быть записана как:

$$p(u_{1,i},u_{2,j}\,|\,h_1)=\int_{\Phi} p(u_{1,i},u_{2,j}\,|\,M)\cdot p(M\,|\,\widehat{U})\cdot dM, \tag{11}$$

где

$p(u_{1,i},u_{2,j}\,|\,M)$ – плотность вероятности измерений $u_{1,i}$ и $u_{2,j}$ при условии, что вектор M отвечает истинному положению соответствующей цели в пространстве признаков.

Плотность $p(u_{1,i},u_{2,j}\,|\,M)$ определяется из условия, что измерения $u_{1,i}$ и $u_{2,j}$ распределены по многомерному нормальному закону с одним и тем же математическим ожиданием M, отвечающим истинному месту цели:

$$p(u_{1,i},u_{2,j}\,|\,M)=\gamma(X_{1,i},M,\Psi_{1,i})\cdot\gamma(X_{2,j},M,\Psi_{2,j})\,. \tag{12}$$

С учетом (8), (9), (10), (11) и (12), выражение для $\mu(u_{1,i},u_{2,j})$ в самом общем случае имеет вид:

$$\mu(u_{1,i},u_{2,j})=\frac{\int\limits_{\Phi}\gamma(X_{1,i},M,\Psi_{1,i})\cdot\gamma(X_{2,j},M,\Psi_{2,j})\cdot p(M\,|\,\widehat{U})\cdot dM}{\int\limits_{\Phi}\int\limits_{\Phi}\dfrac{p(M_2\,|\,M_1,\widehat{U})\cdot p(M_1\,|\,\widehat{U})\cdot dM_2\cdot dM_1}{\gamma(X_{1,i},M_1,\Psi_{1,i})^{-1}\cdot\gamma(X_{2,j},M_2,\Psi_{2,j})^{-1}}}\,. \tag{13}$$

Корректность (13) сохранится при перестановке индексов M_1 и M_2.

Плотности вероятностей $p(M\,|\,\widehat{U})$ и $p(M_2\,|\,M_1,\widehat{U})$ позволяют учесть любой объем априорной информации о пространственных отношениях между целями. Этим обуславливается универсальность РП, получаемого при подстановке (6), (7) и (13) в (5). Однако практическая сложность задания этих плотностей вероятностей затрудняет использование выражения (13) в реальных алгоритмах. Поэтому на основе (13) для конкретных доступных априорных данных необходим синтез более простых, практически применимых выражений.

Синтез частного решающего правила

Наибольшую трудность при отождествлении представляет принятие решений для пар ФЦ, принадлежащих одной цели или двум близкорасположенным целям. Нетождественность ФЦ удаленных целей с высокой достоверностью устанавливается на первом этапе двухэтапной схемы. Поэтому при вычислении $p(u_{1,i},u_{2,j}\,|\,h_0)$ возможно введение предположения о принадлежности этих ФЦ двум близлежащим целям.

Исходя из априорных данных о пространственных отношениях между целями для близлежащих целей возможно определение некоторого характерного расстояния Δb_v ($v=\overline{1,L}$) по каждому v-ому признаку [7]. Кроме того, можно указать минимальные расстояния Δa_v ($\Delta a_v < \Delta b_v$) между целями по каждому признаку, на которые (при равенстве значений других признаков) две цели сближаться не могут. Если эти минимальные расстояния нельзя достоверно определить исходя из априорной информации о пространственных отношениях между целями, они могут быть заданы на основе данных о разрешающих способностях ИИ.

Для определенности, будем считать, что если цель расположена в точке M, то близлежащая цель с постоянной плотностью вероятности находится в одной из точек гиперпараллелограмма с центром в точке M и длиной стороны $2\cdot\Delta b_v$ по v-му признаку, но не лежит в объеме гиперпараллелограмма с центром в точке M и длинами сторон $2\cdot\Delta a_v$. При этом, плотность вероятности нахождения первой цели в каждой точке области пространства признаков, отвечающей ГФЦ, будем считать постоянной и равной p_M. Формально сказанное записывается в виде:

$$p(M_1\,|\,\widehat{U}) = p(M\,|\,\widehat{U}) = p_M\,, \tag{14}$$

$$p(M_2\,|\,M_1,\widehat{U}) = p(M_2\,|\,M_1) = \begin{cases} 0,\; M_2 \notin AB(M_1) \\ p_{AB},\; M_2 \in AB(M_1) \end{cases}, \tag{15}$$

где

$AB(M_1)$ – множество точек гиперпараллелограмма с центром в точке M_1 и длиной стороны $2\cdot\Delta b_v$ по v-ому измерению, не принадлежащих гиперпараллелограмму с центром в точке M_1 и длиной сторон $2\cdot\Delta a_v$ по v-ому измерению;

$$p_{AB} = \frac{1}{2^L\cdot\prod\limits_{v=1}^{L}\Delta b_v - 2^L\cdot\prod\limits_{v=1}^{L}\Delta a_v}\,.$$

Тогда выражение для $\mu(u_{1,i},u_{2,j})$ преобразуется к виду:

$$\mu(u_{1,i},u_{2,j}) = \frac{\int\limits_{\Phi} \gamma(X_{1,i}, M, \Psi_{1,i}) \cdot \gamma(X_{2,j}, M, \Psi_{2,j}) \cdot dM}{\int\limits_{\Phi} \int\limits_{AB(M_1)} \gamma(X_{1,i}, M_1, \Psi_{1,i}) \cdot \gamma(X_{2,j}, M_2, \Psi_{2,j}) \cdot p_{AB} \cdot dM_2 \cdot dM_1} . \tag{16}$$

Числитель (16) выражает плотность вероятности получения разницы $\Delta X_{i,j} = X_{1,i} - X_{2,j}$ по признакам для отождествления при условии, что все элементы вектора математических ожиданий разницы по каждому признаку равны нулям. Поэтому он может быть преобразован к виду [2, 3]:

$$\int\limits_{\Phi} \gamma(X_{1,i}, M, \Psi_{1,i}) \cdot \gamma(X_{2,j}, M, \Psi_{2,j}) \cdot dM = \gamma(\Delta X_{i,j}, O, \Psi^{\Sigma}_{i,j}) , \tag{17}$$

где

$\Psi^{\Sigma}_{i,j} = (\Psi_{1,i} + \Psi_{2,j})$ – суммарная корреляционная матрица ошибок измерения из этих ФЦ;

O – нулевой вектор математических ожиданий.

Аналогично, знаменатель (16), выражая плотность вероятности получения разницы $\Delta X_{i,j} = X_{1,i} - X_{2,j}$ по признакам для отождествления при условии, что вектор математических ожиданий равновероятно принимает значения множества $AB(O)$, может быть преобразован следующим образом:

$$\begin{aligned} &\int\limits_{\Phi} \int\limits_{AB(M_1)} \gamma(X_{1,i}, M_1, \Psi_{1,i}) \cdot \gamma(X_{2,j}, M_2, \Psi_{2,j}) \cdot p_{AB} \cdot dM_2 \cdot dM_1 = \\ &\qquad = p_{AB} \cdot \int\limits_{AB(O)} \gamma(\Delta X_{i,j}, \Delta M, \Psi^{\Sigma}_{i,j}) \cdot d\Delta M \end{aligned} . \tag{18}$$

Подстановкой (17) и (18) в (16) получим:

$$\mu(u_{1,i},u_{2,j}) = \frac{\gamma(\Delta X_{i,j}, O, \Psi^{\Sigma}_{i,j})}{p_{AB} \cdot \int\limits_{AB(O)} \gamma(\Delta X_{i,j}, \Delta M, \Psi^{\Sigma}_{i,j}) \cdot d\Delta M} . \tag{19}$$

При переходах (17) и (18) краевыми эффектами пренебрегаем. Это правомерно, так как область Ф не меньше выпуклой оболочки, построенной на точках, задаваемых значениями признаков из ФЦ.

Рассмотрим случай использования одного признака для отождествления. Тогда после преобразований (19) принимает вид:

$$\mu(u_{1,i},u_{2,j}) = \frac{e^{-\frac{1}{2}\cdot\frac{\Delta x_{i,j}{}^2}{\sigma_{i,j}^2}}\cdot\left(\sqrt{2\cdot\pi}\cdot\sigma_{i,j}\right)^{-1}}{\frac{\Delta b_1\cdot\delta(\Delta x_{i,j},\sigma_{i,j},\Delta b_1)}{(\Delta b_1-\Delta a_1)}-\frac{\Delta a_1\cdot\delta(\Delta x_{i,j},\sigma_{i,j},\Delta a_1)}{(\Delta b_1-\Delta a_1)}}, \tag{20}$$

где

$\Delta x_{i,j} = x_{1,i} - x_{2,j}$ – разность значений признака из ФЦ;

$\sigma_{i,j}^2 = \sigma_{1,i}^2 + \sigma_{2,j}^2$ – суммарная дисперсия ошибок измерения признака для отождествления из ФЦ;

$$\delta(\Delta x,\sigma,\Delta c) = \frac{1}{2\cdot\Delta c}\cdot\int_{-\Delta c}^{\Delta c}\frac{e^{-\frac{1}{2}\cdot\frac{(\Delta x-\Delta m)^2}{\sigma^2}}\cdot d\Delta m}{\sqrt{2\cdot\pi}\cdot\sigma}.$$

Выражение $\delta(\Delta x,\sigma,\Delta c)$ представляет собой функцию плотности вероятности распределения случайной величины, являющейся суммой двух случайных величин:

- распределенной по равномерному закону на интервале $(-\Delta c,\Delta c)$;

- распределенной по нормальному закону с дисперсией σ.

Итоговое распределение хорошо описывается обобщенным распределением Гаусса (ОРГ), функция плотности вероятности которого имеет вид [8]:

$$p(x) = \frac{\lambda}{2\cdot\Gamma(1/\lambda)\cdot w(\lambda)\cdot\sigma_{ONR}}\cdot e^{-\left|\frac{x-m}{w(\lambda)\cdot\sigma_{ONR}}\right|^{\lambda}}, \tag{21}$$

где

$w(\lambda) = \sqrt{\dfrac{\Gamma(1/\lambda)}{\Gamma(3/\lambda)}}$;

m, σ_{ONR} и λ – математическое ожидание, среднеквадратическое отклонение и показатель степени ОРГ соответственно.

Математическое ожидание в данном случае равно нулю, так как является суммой нулевых математических ожиданий слагаемых. Дисперсия суммарной случайной величины является суммой дисперсий слагаемых (их ковариация равна нулю):

$$\sigma_{ONR}^2 = \sigma^2 + \frac{(\Delta c)^2}{3}. \tag{22}$$

Показатель степени λ может быть определен исходя из условия нормировки, которое после несложных преобразований принимает вид:

$$\frac{1}{2 \cdot \Delta c} \cdot \Phi\left(\frac{\Delta c}{\sigma}\right) = \frac{\lambda}{2 \cdot \Gamma(1/\lambda) \cdot w(\lambda) \cdot \sigma_{ONR}}. \tag{23}$$

Аналитическое решение (относительно λ) уравнения (23) представляет собой сложную задачу. Однако стандартной практикой для решения подобных задач в вычислительных системах является использование аппроксимирующих формул. Например, определение $\lambda(\Delta, \sigma)$ с достаточной точностью может осуществляться по формуле:

$$\lambda(\Delta c, \sigma) \approx \left(\frac{e_2}{\frac{\sqrt{\sigma^2 + \Delta c^2 / 3}}{2 \cdot \Delta c} \cdot \Phi\left(\frac{\Delta c}{\sigma}\right) - e_1} \right)^{\frac{1}{e_4}} - e_3, \tag{24}$$

где

$e_1 = 0{,}289$, $e_2 = 1{,}2$, $e_3 = 0{,}8895$, $e_4 = 2{,}2785$ – коэффициенты.

Таким образом, выражение для $\delta(\Delta x, \sigma, \Delta c)$ может быть заменено следующим:

$$\delta'(\Delta x, \sigma, \Delta c) = \frac{\lambda(\Delta c, \sigma) \cdot \exp\left[-\left| \frac{\Delta x}{w(\lambda(\Delta c, \sigma)) \cdot \sqrt{\sigma^2 + \frac{(\Delta c)^2}{3}}} \right|^{\lambda(\Delta c, \sigma)} \right]}{2 \cdot \Gamma\left(\frac{1}{\lambda(\Delta c, \sigma)}\right) \cdot w(\lambda(\Delta c, \sigma)) \cdot \sqrt{\sigma^2 + \frac{(\Delta c)^2}{3}}}. \tag{25}$$

При переходе к использованию L признаков для отождествления можно учесть, что знаменатель (19) определяет функцию плотности вероятности многомерного (при переходе к L признакам) нормального распределения с недиагональной в общем случае корреляционной матрицей (ошибки ИИ), «разбавленного» многомерным распределением (пространственные отношения между целями) без корреляции. Это дает возможность пренебречь корреляцией в полученном распределении и записать выражение для $\mu(u_{1,i}, u_{2,j})$ в следующем виде:

$$\mu(u_{1,i}, u_{2,j}) = \frac{\gamma(\Delta X_{i,j}, O, \Psi_{i,j}^{\Sigma})}{\dfrac{\prod_{\nu=1}^{L} \delta'(\Delta x_{i,j,\nu}, \sigma_{i,j,\nu}, \Delta b_\nu)}{V_{\Delta b}^{-1} \cdot (V_{\Delta b} - V_{\Delta a})} - \dfrac{\prod_{\nu=1}^{L} \delta'(\Delta x_{i,j,\nu}, \sigma_{i,j,\nu}, \Delta a_\nu)}{V_{\Delta a}^{-1} \cdot (V_{\Delta b} - V_{\Delta a})}}, \tag{26}$$

где

$V_{\Delta b} = 2^L \cdot \prod_{\nu=1}^{L} \Delta b_\nu$ ($V_{\Delta a} = 2^L \cdot \prod_{\nu=1}^{L} \Delta a_\nu$) – объем L-мерного гиперпараллелограмма с длинной стороны по ν-ой оси $2 \cdot \Delta b_\nu$ ($2 \cdot \Delta a_\nu$);

$\Delta x_{i,j,\nu}$ – ν-ый элемент вектора $\Delta X_{i,j}$;

$\sigma_{i,j,\nu}^2$ – ν-ый диагональный элемент матрицы $\Psi_{i,j}^{\Sigma}$.

При подстановке (6), (7) и (26) в (5) получается частное РП, использующее практически доступные объемы априорных данных о пространственных отношениях между целями. Нужно также отметить, что при невозможности указания величин Δb_ν они могут быть положены стремящимися (близкими) к Δa_ν ($\Delta b_\nu \to \Delta a_\nu$). Такое предположение будет негативно сказываться на ВПР, однако, даст возможность принимать решения в условиях полного отсутствия априорных данных о пространственных отношениях между целями (при выборе Δa_ν на основе разрешающих способностей ИИ).

Синтез алгоритма

Признак для сравнения ВО в оптимальном по критерию максимума апостериорной вероятности алгоритме отождествления внутри ГФЦ может быть получен на основе (5) (с учетом (6) и (7)) путем логарифмирования, отбрасывания постоянных слагаемых и перегруппировки:

$$\alpha_1(A_g) = \ln\left(\frac{(N_1+N_2-k_g)!\cdot(q_2\cdot q_1)^{N_1+N_2-k_g}}{(1-q_2\cdot q_1)^{N_1+N_2-k_g+1}}\right) + \sum_{m=1}^{k_g}\ln\mu(u_{1,y_{1,m}^g}, u_{2,y_{2,m}^g})\,. \tag{27}$$

Для синтеза алгоритма аппроксимируем первое слагаемое линейной функцией и, после отбрасывания постоянных членов, получим:

$$\alpha_2(A_g) = \sum_{m=1}^{k_g}(a + \ln\mu(u_{1,y_{1,m}^g}, u_{2,y_{2,m}^g}))\,, \tag{28}$$

где

$a = \dfrac{6\cdot\sum_{i=0}^{N}(2\cdot i - N)\cdot f_1(i)}{N^3 + 3\cdot N^2 + 2\cdot N}$ – угловой коэффициент аппроксимирующей прямой, определенный по методу наименьших квадратов;

$f_1(k) = \sum_{i=1}^{N_1+N_2-k}\ln i + (N_1+N_2-k)\cdot\ln q_1\cdot q_2 - (N_1+N_2-k+1)\cdot\ln(1-q_2\cdot q_1)$ – логарифм априорной вероятности (без постоянных слагаемых);

$N = \min(N_1, N_2)$.

Как показывают расчеты, выполненные для большого числа частных случаев, аппроксимация, введенная в (28), обычно очень точна [5].

Рассмотрим двудольный мультиграф $G_d(V_d, E_d)$, множества вершин V_d^1 и V_d^2 ($V_d^1 \cup V_d^2 = V_d$; $V_d^1 \cap V_d^2 = \{\}$) долей которого соответствуют множествам ФЦ от двух ИИ в ГФЦ, а между любой парой вершин $\{i, j\}$ разных долей ($i \in V_d^1$; $j \in V_d^2$) в множестве ребер E_d существует два кратных ребра с весами, определяемыми по формулам:

$$r_{i,j}^1 = a + \ln\mu(u_{1,i}, u_{2,j})\,, \tag{29}$$

$$r_{i,j}^2 = 0\,. \tag{30}$$

Выберем в этом мультиграфе произвольное совершенное паросочетание (СП) e, включающее D_e^1 ребер первого и D_e^2 ребер второго типа. По определению СП, $D_e^1 + D_e^2 = \min(N_1, N_2)$. Перенумеровав элементы e так, чтобы ребра первого типа имели номера с 1 по D_e^1, а ребра второго типа – с $(D_e^1 + 1)$ по $(D_e^1 + D_e^2)$, сумму весов e можно записать как:

$$S(e) = \sum_{w=D_e^1+1}^{D_e^1+D_e^2} r_{e_w}^2 + \sum_{w=1}^{D_e^1} r_{e_w}^1 = \sum_{w=1}^{D_e^1} (a + \ln \mu(u_{1,i_w}, u_{2,j_w})), \tag{31}$$

где

$e_w = \{i_w, j_w\}$ – пара вершин, инцидентных w-му ребру из СП e.

Вес произвольного СП e в $G_d(V_d, E_d)$ дает значение признака (28) для ВО A_g, в соответствии с которым отождествляется $D_e^1 = k_g$ пар ФЦ, отвечающих ребрам первого типа из e. В силу произвольности выбора e, любое СП однозначно соответствует некоторому ВО. Легко показать корректность и обратного утверждения: каждому ВО однозначно соответствует не менее одного СП в мультиграфе $G_d(V_d, E_d)$.

Таким образом, задача сводится к поиску СП максимального веса (СПМВ) в двудольном мультиграфе $G_d(V_d, E_d)$. Для ее решения мультиграф $G_d(V_d, E_d)$ может быть преобразован в простой двудольный граф $G(V_d, E)$ исключением из каждой пары кратных ребер минимального по весу ребра. В полученном графе между любой парой вершин $\{i, j\}$ разных долей будет существовать ребро с весом:

$$r_{i,j} = \max\left(r_{i,j}^1, r_{i,j}^2\right). \tag{32}$$

СПМВ e' в графе $G(V_d, E)$ будет содержать ребра, инцидентные парам вершин, ребра между которыми входят в СПМВ в $G_d(V_d, E_d)$.

В ходе преобразования $G_d(V_d, E_d)$ в $G(V_d, E)$ исключаются решения для пар ФЦ, заведомо не входящие в искомый ВО. Множество пар номеров ФЦ в ГФЦ, которые должны быть отождествлены в соответствии с ВО, максимизирующим признак (28), на основе СПМВ e' в графе $G(V_d, E)$ определяется следующим образом:

$$Q = \{\{i, j\} \mid [\{i, j\} \in e'] \wedge [a + \ln \mu(u_{1,i}, u_{2,j}) > 0]\}, \quad (33)$$

где

i, j – номера отождествляемых ФЦ от первого и второго ИИ.

Задача поиска СПМВ в простом полном двудольном графе хорошо известна и существуют алгоритмы ее решения, обладающие сложностью $O(n^3)$, где n – количество вершин в большей доле графа [1].

Рассмотренный алгоритм отождествления внутри ГФЦ должен быть применен к каждой выделенной ГФЦ. Сложность этого алгоритма и, соответственно, всего алгоритма отождествления данных от двух ИИ в общем случае составляет $O(n^3 + n^2 \cdot \chi(\cdot))$, где $\chi(\cdot)$ – сложность нахождения значения логарифма $\mu(u_{1,i}, u_{2,j})$. При использовании частных РП (например, рассчитывающего $\mu(u_{1,i}, u_{2,j})$ по формуле (26)) $\chi(\cdot) = const$ и сложность предложенного алгоритма становится полиномиальной $O(n^3)$, где n – число ФЦ от одного ИИ.

Результаты экспериментального исследования

Оценка адекватности предложенного подхода проводилась посредством математического моделирования. Для отождествления использовались пеленга и дистанции 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 СКО определения соответствующего признака ИИ, d – коэффициент. При моделировании задавались различные значения других параметров (количество целей, вероятности взятия целей на сопровождение ИИ и др.). Отношение плотностей вероятности $\mu(u_{1,i}, u_{2,j})$ определялось по формуле (26) (использовалось частное РП). Априорное минимальное расстояние между целями по каждому признаку (при условии равенства значений других признаков) задавалось на основе соотношения между СКО и разрешающей способностью ИИ – $\Delta a_v = 1{,}1 \cdot \sigma_v$. Характерные расстояния между целями по признакам задавались равными $\Delta b_v = 2{,}0 \cdot d \cdot \sigma_v$.

По результатам моделирования было установлено, что высокие ВПР (выше 0,75..0,80) разработанным алгоритмом достигаются при расстояниях между целями хотя бы по одному признаку, превышающих в 1,6..1,9 раза суммарную СКО определения соответствующего признака ИИ.

Кроме того, было проведено сравнение ВПР предложенным алгоритмом и алгоритмами, описанными в [3, 4, 5, 6, 7], которое показало, что разработанный алгоритм в большинстве ситуаций обеспечивает более высокие вероятностные характеристики. Величина прироста ВПР при использовании рассмотренного подхода зависит от конкретных условий, в первую очередь, от вероятностей взятия целей на сопровождение.

Заключение

В статье предложены оптимальные по критерию максимума апостериорной вероятности универсальное и частное РП для отождествления внутри ГФЦ. Эти РП отличаются от известных обеспечением более высоких вероятностей принятия правильных решений, что достигнуто снятием ряда допущений (гипотез). Можно показать, что многие известные РП (например, описанные в [3, 4, 5, 7]) являются частными случаями предложенных.

На основе синтезированных РП разработан алгоритм, при использовании частного РП обладающий полиномиальным временем работы. Данный алгоритм составляет основу для решения задачи отождествления данных от двух ИИ и, следовательно, от произвольного числа ИИ.

Малая вычислительная сложность и более высокие вероятностные характеристики обуславливают возможность применения рассмотренных процедур при создании алгоритмического и программного обеспечения решения задачи отождествления при модернизации существующих и создании новых АИУС.

Библиографический список

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

2. Ширман Я.Д., Манжос В.Н. Теория и техника обработки радиолокационной информации на фоне помех – М: Радио и связь, 1981. – 416 с.

3. Кузьмин С.З. Основы теории цифровой обработки радиолокационной информации – М.: Сов. радио, 1974. – 432 с.

4. Музыченко О.Н., Савченко Д.И. Использование теории графов для отождествления целевой информации // Сборник докладов научно-технической конференции «Состояние,

проблемы и перспективы создания корабельных информационно-управляющих комплексов». – М: ОАО «Концерн «Моринформсистема-Агат», 2009. – С. 132-136.

5. Савченко Д.И. Задача отождествления данных в системах освещения обстановки // Итоги диссертационных исследований. Том 3. – Материалы III Всероссийского конкурса молодых ученых. – М.:РАН, 2011. – С. 149-159.

6. Савченко Д.И. Эвристические алгоритмы для второго этапа отождествления информации // Сборник докладов научно-технической конференции «Состояние, проблемы и перспективы создания корабельных информационно-управляющих комплексов (эффективность, надежность, экономика)». – М.: ОАО «Концерн «Моринформсистема-Агат», 2011. – С. 42-47.

7. Автоматизация обработки, передачи и отображения радиолокационной информации / Под ред. В.Г. Корякова. – М.: Сов. Радио, 1975. – 304 с.

8. Булашев С.В. Статистика для трейдеров. – М.: Компания Спутник+, 2003. – 245 с.

Сведения об авторе

Савченко Даниил Игоревич, инженер-программист I категории ОАО «Научно-производственная фирма «Меридиан»;
тел.: +7 (911) 286-43-99, e-mail: cfdX@list.ru.

Работа выполнена при поддержке Минобрнауки России в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 годы».