УДК 519.85

МОДЕЛЬ И МЕТОД РЕШЕНИЯ ЗАДАЧИ ДИСКРИМИНАЦИИ ТРАЕКТОРИЙ ПРИ СКАНИРОВАНИИ БИОЛОГИЧЕСКИХ ОБЪЕКТОВ

к.т.н. Ю.Е. Мегель
(представил д.т.н., проф. В.П. Путятин)

В статье представлена модель и метод решения задачи дискриминации типов и параметров траекторий биологических объектов, полученных при решении задачи идентификации траекторий по результатам сканирования.

1. Постановка задачи. На плоскости с системой координат $\{O;x;y\}$ дана выпуклая область F. В начальный момент времени $j=0$ в ней имеется n объектов $I=\{1, 2, ..., n\}$ с координатами $\{(x_i^0, y_i^0)\}_{i=1,2,...,n}$. В последующем объекты могут перемещаться по плоскости, выходя за пределы области F. По качественному характеру своего движения объекты относятся к одному из трех классов: K_1 – неподвижные; K_2 – манежные (перемещаются по окружности); K_3 – линейные (перемещаются по прямой или параболе).

В течение промежутка времени T объекты могут перемещаться со скоростью, не превосходящей v_{max}. В этот период их координаты сканируются в дискретные моменты времени $j=1, 2, ..., m$ с шагом $\tau = T/m$.

За период T большая часть объектов не покидает область F. Объект, покинувший ее, может вернуться. На шаге сканирования j фиксируются координаты объектов в области F без их соотнесения с номерами объектов, идентифицированных на шаге $j-1$. Если два объекта сблизились на расстояние $\delta \le 2\tau v_{max}$, то они могут взаимодействовать с изменением направления движения. Многократное взаимодействие практически невозможно.

Для этого процесса требуется разработать метод детерминированной и статистической дискриминации объектов I по классам K_1, K_2, K_3 в реальном масштабе времени, включающий оценку энергетических параметров объектов: p_l – долю объектов класса K_l, $(l=1, 2, 3)$; $\bar{r}_0$ – радиус биений для объектов класса K_1; $\bar{r}_M$ – средний радиус обращения для объектов класса K_2; $\bar{V}$ – скорость перемещения объектов класса.

2. Модель и метод идентификации траекторий. Пусть $\mathbf{Z_j} = \{\mathbf{z_{\eta\, j} = (x_{\eta\, j}, y_{\eta\, j})}\}_{\eta \in \mathbf{I_J}}$ – множество точек в области $\mathbf{F}$, координаты которых получены на шаге $\mathbf{j \in J}$; назовем их субъектами. На начальном шаге $\mathbf{(j = 0)}$ пронумеруем эти точки в произвольном порядке и назовем их объектами. Положим $\mathbf{I = I_0 = \{1, 2, \ldots, n\}}$. Объект $\boldsymbol{\mu}$, вошедший в $\mathbf{F}$ на шаге $\mathbf{j \geq 1}$, назовем внешним и положим $\boldsymbol{\mu} \in \mathbf{I^* = \{I \cup \{n+1, n+2, \ldots\}\}}$, $\boldsymbol{\mu} \geq \mathbf{n+1}$. На каждом последующем шаге $\mathbf{j}$, $\mathbf{(j \geq 1)}$, необходимо решать задачу идентификации субъектов $\{\mathbf{z_{\eta j}}\}_{\eta \in \mathbf{I_J}}$ – присваивать каждому субъекту $\boldsymbol{\eta} \in \mathbf{I_j}$ номер $\boldsymbol{\xi} \in \mathbf{I^*}$ того объекта, который на шаге $\mathbf{j}$ перешел в точку $\mathbf{z_{\eta j}}$. При этом возможны следующие ситуации: {1} – объект $\boldsymbol{\xi}$ остался в $\mathbf{F}$ и не имел взаимодействий; {2} – объект $\boldsymbol{\xi}$ покинул $\mathbf{F}$; {3} – объект $\boldsymbol{\xi}$ остался в $\mathbf{F}$ и имел взаимодействие; {4} – некоторый внешний объект $\boldsymbol{\mu}$ вошел в $\mathbf{F}$. Их обработка производится следующим образом.

Ситуация {1}. Идентификация основана на том, что при отсутствии взаимодействий объекта $\boldsymbol{\xi}$ на шаге $\mathbf{j}$, в $\boldsymbol{\varepsilon}_\mathbf{m}$ - окрестности $\boldsymbol{\varepsilon}_{\boldsymbol{\xi},\mathbf{j-1}}$ точки $\mathbf{s}_{\boldsymbol{\xi},\mathbf{j-1}}$, определяемой кругом радиуса $\boldsymbol{\varepsilon}_\mathbf{m} = \mathbf{v_m}\boldsymbol{\tau}$ с центром в точке $\mathbf{s}_{\boldsymbol{\xi},\mathbf{j-1}}$, может содержаться лишь субъект $\mathbf{z_{\eta j}}$, определяющий положение объекта $\boldsymbol{\xi} \in \mathbf{I}$ на шаге $\mathbf{j}$. Соответствующая последовательность точек $\mathbf{J_{\xi\, j} = (x_{\xi l}, y_{\xi l})\}_{l=0,1,\ldots,j}}$ определяет траекторию объекта $\boldsymbol{\xi}$ на интервале $\mathbf{T' = j\tau}$.

Ситуация {2}. Объекту $\boldsymbol{\xi}$ не соответствует ни один субъект $\boldsymbol{\eta} \in \mathbf{I_J}$. Поэтому выводим этот объект из рассмотрения при отслеживании траекторий, усеченную траекторию $\mathbf{J^*_{\xi j} = \{s_{\xi l} = (x_{\xi l}, y_{\xi l})\}_{l=0,\ldots,j-1}}$ для объекта $\boldsymbol{\xi} \in \mathbf{I}$ запоминаем, а траекторию для объекта $\boldsymbol{\mu} \in \mathbf{I^* \setminus I}$ отбрасываем.

Ситуация {3}. Имеем взаимодействие двух объектов. Это означает, что на шаге $\mathbf{j-1}$ некоторая пара объектов $\boldsymbol{\xi}, \boldsymbol{\eta}$ имеет пересекающиеся $\boldsymbol{\varepsilon}_\mathbf{m}$-окрестности $\boldsymbol{\varepsilon}_{\boldsymbol{\xi},\mathbf{j-1}}$, $\boldsymbol{\varepsilon}_{\boldsymbol{\eta},\mathbf{j-1}}$. Поскольку после взаимодействия ориентация траекторий может быть изменена, считаем траектории $\mathbf{J^*_{\xi j} = \{s_{\xi l}\}_{l=0,\ldots,j-1}}$ и $\mathbf{J^*_{\eta j} = \{s_{\eta l}\}_{l=0,\ldots,j-1}}$ усеченными и начинаем отслеживать для этих объектов новые траектории $\mathbf{J'_{\xi k} = \{s_{\xi l}\}_{l=j,\ldots,k}}$, $\mathbf{J'_{\eta k} = \{s_{\eta l}\}_{l=j,\ldots,k}}$, $\mathbf{k > j}$.

Ситуация {4}. Присваиваем объекту очередной номер и заносим его в $\mathbf{I}^* = \{\mathbf{I} \cup \{\mathbf{n+1, n+2, ...}\}\}$; эти объекты отслеживаются (по двум шагам), но не дискриминируются.

В итоге, по окончании процесса сканирования для каждого исходного объекта $\xi \in \mathbf{I}$ имеем одну полную траекторию $\mathbf{J}_{\xi \mathbf{m}}$, если объект ξ не имел взаимодействий и не покинул область $\mathbf{F}$; либо одну усеченную траекторию $\mathbf{J}^*_{\xi \mathbf{k}}$, $\mathbf{k < m}$, если объект ξ покинул область $\mathbf{F}$ на шаге $\mathbf{k}$; либо две усеченных траектории – $\mathbf{J}^*_{\xi \mathbf{k}}$ (до взаимодействия на шаге $\mathbf{k}$) и $\mathbf{J}'_{\xi \mathbf{k}}$ или $\mathbf{J}^{(1)}_{\xi\eta,\, \mathbf{j}}$ (движение объекта после этого в области $\mathbf{F}$ или до выхода из нее).

3. Метод дискриминации типов траекторий. В целях упрощения индексации рассмотрим объект $\mathbf{i} \in \mathbf{I}$, который не имел взаимодействий и не покинул области $\mathbf{F}$ с траекторией $\mathbf{J_{i\,m}} = \{\mathbf{s_{il}} = (\mathbf{x_{il}}, \mathbf{y_{il}})\}_{\mathbf{l=0,1,...,m}}$.

Дискриминация класса K_1. Если $\mathbf{i} \in \mathbf{K_1}$ точки $\mathbf{J_{i\,m}}$, то в среднем должны располагаться в круге достаточно малого радиуса $\mathbf{R_0}$, в котором объект $\mathbf{i}$ совершает флуктуации, обусловленные броуновским движением и собственной малой подвижностью. Пусть $\mathbf{z_i} = (\mathbf{x_z}, \mathbf{y_z})$ – центр тяжести точек $\mathbf{J_{i\,m}}$; $\{\mathbf{r_l}\}_{\mathbf{l=0,m}}$ – расстояния от них до точки $\mathbf{z_i}$, а $\overline{\mathbf{R}}$, $\mathbf{S}$ – их среднее и СКО. Введем следующие критерии для дискриминации объекта $\mathbf{i}$ к классу $\mathbf{K_1}$:

$$\mathbf{f_1^{(1)}}: \quad \max_{\mathbf{l=0,1,...,m}} [(\mathbf{x_{il}} - \mathbf{x_z})^2 + (\mathbf{y_{il}} - \mathbf{y_z})^2]^{1/2} \le \mathbf{R_0}\ ; \tag{1}$$

$$\mathbf{f_2^{(1)}}: \quad \overline{\mathbf{R}} \le \mathbf{R_0}; \tag{2}$$

$$\mathbf{f_3^{(1)}}: \quad \mathbf{P[H : r < R_0] \ge \beta}. \tag{3}$$

Критерий (1) определяет жесткую детерминацию условия о том, что объект $\mathbf{i}$ не удалился от начальной точки $(\mathbf{x_i^0}, \mathbf{y_i^0})$ на расстояние, превышающее $\mathbf{2R_0}$, а в среднем – совершает флуктуации в пределах круга радиуса $\mathbf{R_0}$. Он полезен при выборках малого объема и там, где следует выделить максимальное число объектов, обладающих малой подвижностью.

Критерий (2) определяет, что средняя флуктуация объектов не превышает значения $\mathbf{R_0}$, характерного для данных условий эксперимента (по температуре, плотности анализируемой среды и др.). В смысле жесткости дискриминации он занимает промежуточное положение между (5) и (7).

Критерий (3) определяет, что с доверительной вероятностью β истинное отклонение $\mathbf{r}$ объекта от центра флуктуаций не превышает предельного значения $\mathbf{R_0}$. Этот критерий полезен там, где нет возможности задаться точным значением. В силу независимости блужданий объекта и идентичности типа его движений по шагам. В первом приближении можно считать, что среднее $\overline{\mathbf{R}}$ стремится к нормальному распределению $\mathbf{N(z_i, S^2)}$. Тогда на уровне значимости $\alpha = 1 - \beta$ можем принять гипотезу (3) о принадлежности объекта $\mathbf{i}$ классу $\mathbf{K_1}$ при выполнении неравенства

$$\frac{\overline{\mathbf{R}} - \mathbf{R_0}}{\mathbf{S}}\sqrt{\mathbf{m+1}} \le \mathbf{u}_\alpha \text{ при } \mathbf{m > 50}, \left(\frac{\overline{\mathbf{R}} - \mathbf{R_0}}{\mathbf{S}}\sqrt{\mathbf{m}} \le \mathbf{t}_{\alpha/2;\mathbf{m}} \quad \text{при} \quad \mathbf{m \le 50} \right), \tag{4}$$

где $\mathbf{u}_\alpha$ ($\mathbf{t}_{\alpha/2,\mathbf{m}}$) - квантиль нормального (Стьюдента) распределения.

Дискриминация класса К₂. Если объект $\mathbf{i}$ принадлежит классу $\mathbf{K_2}$, то его траектория $\mathbf{J_{i\,m}}$ в силу возмущений и собственных флуктуаций не является окружностью и, в среднем, должна располагаться в пределах кольца с радиусами $\mathbf{R_1}$ и $\mathbf{R_2}$, $\mathbf{R_0 \le R_1 < R_2}$, и центром в соответствующей точке $\mathbf{z_i = (x_z, y_z)}$. Поэтому для дискриминации манежного объекта достаточно установить, что радиус его орбиты лежит в соответствующих пределах. Для этого используем обобщения критериев (1) – (3) на случай двух неравенств.

Дискриминация класса К₃. Если объект $\mathbf{i}$ не относится к классам $\mathbf{K_1}$ или $\mathbf{K_2}$, то его можно просто отнести к классу $\mathbf{K_3}$, либо воспользоваться известной моделью линейной регрессии, которая позволяет с достаточной точностью отличить траекторию, близкую к прямой или параболе, от замкнутой траектории манежного типа. Для этого достаточно сравнить соответствующее отношение регрессионных дисперсий с критической точкой $\mathbf{F_{\alpha, f_1, f_2}}$ распределения Фишера для уровня значимости $\alpha = 1 - \beta$ и числа степеней свободы $\mathbf{f_1 = 1, f_2 = m - 1}$. При этом более точная оценка скорости объекта, не связанная с использованием линейной модели, определяется отношением фактического перемещения и времени $\mathbf{T}$:

$$\overline{\mathbf{V}} = \frac{\mathbf{1}}{\mathbf{T}}\sum_{\mathbf{l=1}}^{\mathbf{m}}[(\mathbf{x_{i,l-1}} - \mathbf{x_{i,l}})^2 + (\mathbf{y_{i,l-1}} - \mathbf{y_{i,l}})^2]^{1/2}\,. \tag{5}$$

4. Оценка трудоемкости и особенности применения метода. Данный метод включает две процедуры: идентификацию объектов после каждого шага сканирования и дискриминацию траекторий по типам по завершению всего процесса сканирования. Оценим его вычислительную трудоемкость, выраженную числом арифметических операций. Трудоемкость идентификации субъектов на одном шаге сканирования составит величину порядка

$$\kappa_s \approx 6 \cdot \mathbf{n}^2, \tag{6}$$

а применение критериев (1) - (3) – оценивается величинами:

$$\kappa_1^{(1)} \approx 7m\,; \qquad \kappa_2^{(1)} \approx 8m\,; \qquad \kappa_1^{(1)} \approx 10m\,. \qquad (7)$$

Трудоемкость же расчета скорости линейного объекта и проверки гипотезы о не манежном характере движения можно оценить величинами порядка:

$$\kappa_V \approx 6m\,; \qquad \kappa_L \approx 15m\,. \qquad (8)$$

Тогда в предположении о том, что объекты пропорционально распределены по классам, средняя трудоемкость идентификации одной траектории соответственно критериям типа (1) – (3) составит:

$$\kappa_1^* \approx 10m\,; \quad \kappa_2^* \approx 10m\,; \quad \kappa_2^* \approx 12m\,. \qquad (9)$$

Проведение регрессионного анализа увеличивает эти оценки на величину $\kappa_L^* = \frac{1}{3} \cdot \kappa_L = 5m.$ Тогда верхняя оценка трудоемкости дискриминации траекторий всех объектов по типу и параметрам составит величину порядка

$$\kappa_I \approx c \cdot m \cdot n\,, \qquad (10)$$

где c принимает значение в диапазоне от 10 – 12 до 15 – 17 в зависимости от выбора критерия (1) – (3) и стратегии анализа линейных траекторий.

Поскольку процедура идентификации на каждом шаге сканирования должна выполняться в реальном масштабе времени, производительность λ^* системы должна удовлетворять неравенству $\lambda^* \geq \kappa_s / \tau$. Так как дискриминация траекторий производится после завершения всего процесса сканирования, то она не критична для обеспечения работы оператора по перезагрузке исследуемых образцов. Исходя из линейной (10) зависимости трудоемкости κ_I от размерности задачи дискриминации, время ее решения незначимо по сравнению с реакцией оператора.

5. Выводы. Предложенный метод оценивания энергетических характеристик биологически активных жидкостей по их кинематическим параметрам может быть применен в режиме реального времени (как описано выше) и апостериорно (в случае запоминания данных обо всех шагах сканирования). При этом после дискриминации объектов по классам и оценивания их параметров по приведенным выше моделям, расчет требуемых характеристик образца – n_l (p_l), $\bar{r}_0$, $\bar{r}_M$, $\overline{V}$ и СКО для них производится стандартным образом на основе статистических методов точечного и интервального оценивания.

Поступила в редколлегию 29.01.2001