

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 $I = \{1, 2, \dots, n\}$ с координатами $\{(x_i^0, y_i^0)\}_{i=1,2,\dots,n}$. 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 – маневренные (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 $j = 1, 2, \dots, m$ 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 сблизались на расстояние $\delta \leq 2\tau v_{\max}$, 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_1 – долю объектов класса K_1 , ($i = 1, 2, 3$); \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 $Z_j = \{z_{\eta j} = (x_{\eta j}, y_{\eta j})\}_{\eta \in I_j}$ – множество точек в области F , координаты которых получены на шаге $j \in J$; назовем их субъектами. На начальном шаге ($j=0$) пронумеруем эти точки в произвольном порядке и назовем их объектами. Положим $I = I_0 = \{1, 2, \dots, n\}$. Объект μ , вошедший в F на шаге $j \geq 1$, назовем внешним и положим $\mu \in I^* = \{I \cup \{n+1, n+2, \dots\}\}$, $\mu \geq n+1$. На каждом последующем шаге j , ($j \geq 1$), необходимо решать задачу идентификации субъектов $\{z_{\eta j}\}_{\eta \in I_j}$ – присваивать каждому субъекту $\eta \in I_j$ номер $\xi \in I^*$ того объекта, который на шаге j перешел в точку $z_{\eta j}$. При этом возможны следующие ситуации: {1} – объект ξ остался в F и не имел взаимодействий; {2} – объект ξ покинул F ; {3} – объект ξ остался в F и имел взаимодействие; {4} – некоторый внешний объект μ вошел в F . Их обработка производится следующим образом.

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

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

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

Ситуация {4}. Присваиваем объекту очередной номер и заносим его

в $\Gamma^* = \{I \cup \{n+1, n+2, \dots\}\}$; эти объекты отслеживаются (по двум шагам), но не дискриминируются.

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

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

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

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

$$f_2^{(1)}: \bar{R} \leq R_0; \quad (2)$$

$$f_3^{(1)}: P\{N: r < R_0\} \geq \beta. \quad (3)$$

Критерий (1) определяет жесткую детерминацию условия о том, что объект i не удалился от начальной точки (x_i^0, y_i^0) на расстояние, превышающее $2R_0$, а в среднем – совершает флуктуации в пределах круга радиуса 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 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 Γ объекта от центра флуктуаций не превышает предельно-

го 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 $\bar{\mathbf{R}}$ стремится к нормальному распределению $\mathbf{N}(\mathbf{z}_i, \mathbf{S}^2)$. Тогда на уровне значимости $\alpha = 1 - \beta$ можем принять гипотезу (3) о принадлежности объекта \mathbf{i} классу \mathbf{K}_1 при выполнении неравенства

$$\frac{\bar{\mathbf{R}} - \mathbf{R}_0}{\mathbf{S}} \sqrt{m+1} \leq \mathbf{u}_\alpha \text{ при } m > 50, \left(\frac{\bar{\mathbf{R}} - \mathbf{R}_0}{\mathbf{S}} \sqrt{m} \leq \mathbf{t}_{\alpha/2; m} \text{ при } m \leq 50 \right), \quad (4)$$

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

Дискриминация класса \mathbf{K}_2 . 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 , $\mathbf{R}_0 \leq \mathbf{R}_1 < \mathbf{R}_2$, 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 \mathbf{K}_3 . 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 $f_1 = 1, f_2 = m - 1$. При этом более точная оценка скорости объекта, не связанная с использованием линейной модели, определяется отношением фактического перемещения и времени \mathbf{T} :

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

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

$$\mathbf{k}_s \approx 6 \cdot n^2, \quad (6)$$

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

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

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

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

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

$$\kappa_1^* \approx 10m; \quad \kappa_2^* \approx 10m; \quad \kappa_3^* \approx 12m. \quad (9)$$

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

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

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_I (p_I), \bar{r}_0 , \bar{r}_M , \bar{V} и СКО для них производится стандартным образом на основе статистических методов точечного и интервального оценивания.

Поступила в редколлегию 29.01.2001