УДК 004.432.4

А. А. ШАРИПБАЕВ, В. Н. РИФА

СТАТИСТИЧЕСКИЕ СВОЙСТВА ТРАЕКТОРИЙ КУРСОРА В ПРОСТРАНСТВЕ ХАРАКТЕРИСТИК

Рассматривается возможность применения метода динамических характеристик для решения задачи идентификации пользователя компьютера по управлению манипулятором «мышь». Численные эксперименты показывают сходимость собственных чисел ковариационных матриц к устойчивому индивидуальному набору для пользователя с ростом числа наблюдений.

Введение. В последнее время заметно возросло количество публикаций по биометрическим методам идентификации личности или биометрическим системам контроля доступа (БСКД). О преимуществах и недостатках таких систем написано очень много. Поэтому не будем повторять известные всем сравнения и остановим свое внимание только на динамических БСКД.

Наиболее просто получить биометрические данные пользователя компьютера можно через анализ управления манипулятором «мышь» тем же пользователем. Рассмотрим целенаправленные движения курсора под управлением пользователя компьютера. Авторам представляется, что именно целенаправленные движения несут более полную информацию о психофизических особенностях субъекта, нежели случайные блуждания курсора, что собственно и было подтверждено результатами многочисленных экспериментов.

В имеющейся литературе неоднократно встречаются попытки использовать нейронные сети (НС) совместно с параметрическими методами для создания классификатора решения задачи динамической БСКД через управление манипулятором. При этом надо учесть тот факт, что при дополнении еще одного пользователя к базе уже существующих, необходимо переобучение НС. Поскольку НС определяет принадлежность входных векторов измерений только тем классам, на которых происходило обучение, обучить НС на класс “всех остальных” не представляется возможным [1, 2].

Естественно, возникает потребность и желание найти метод, который позволил бы строить классификатор не на относительных различиях между классами-субъектами, будь то параметрические или непараметрические статистики, а на некоторых абсолютных значениях вектора признаков. Эти абсолютные значения, или некоторые аттракторы в признаковом пространстве, неоспоримо существуют для каждого субъекта и являются присущими только ему, хотя и существование «двойников» вовсе не исключается. Последнее относится к задаче чувствительности метода, на основе которого строится признаковое пространство и классификатор.

Постановка задачи. Ставится задача распознавания пользователя по множеству статистических параметров, полученных на основе данных движения курсора пользователя в реальном времени. Измерения координат движения курсора необходимо преобразовать таким образом, чтобы получить множество признаков, однозначно определяющих выбранного пользователя системы.

Решение задачи. В работах [4, 9-13] предложен метод построения признакового пространства Ω на основе множества $\Phi(l)$ – характеристик траекторий движения курсора под управлением субъекта. Приведены статистические доказательства существования аттракторов, в пространстве Λ – спектров ковариационных матриц Σ_Φ, индивидуальных для каждого субъекта.

Необходимо найти ответ на вопрос, каким же образом можно простроить классификатор на таких аттракторах? Попытки применить известные теоремы о среднем и использовать классические метрики вида: Если ото-

$r = \left(\left|x_1 - \overline{x}_1\right|^p + \ldots + \left|x_n - \overline{x}_n\right|^p\right)^{\frac{1}{p}}$, хотя и дают приемлемые результаты, однако требуют больших объемов измерений и вычислений, в силу упоминаемых в тех же работах *фрактальных свойств* рядов собственных чисел. Известно, что среднее не является устойчивым для величин временного ряда, имеющего фрактальную природу.

Признаковое пространство на траекториях $\{l\}$ движения курсора определено следующим образом [4]:

$\Phi(l) = \{\varphi_k(x(t_i), y(t_i))\}$, $i = 0,1,\ldots n$, $k = 1,\ldots,6$,

где в качестве характеристик траекторий выбраны следующие функции:

$$vx_i = vx(t_i) = \frac{x_{i+1} - x_{i-1}}{t_{i+1} - t_{i-1}} \quad i = 1,\ldots,n-1;$$

$$vy_i = vy(t_i) = \frac{y_{i+1} - y_{i-1}}{t_{i+1} - t_{i-1}} \quad i = 1,\ldots,n-1;$$

$$ax_i = ax(t_i) = \frac{vx_{i+1} - vx_{i-1}}{t_{i+1} - t_{i-1}} \quad i = 2,\ldots,n-2;$$

$$ay_i = ay(t_i) = \frac{vy_{i+1} - vy_{i-1}}{t_{i+1} - t_{i-1}} \quad i = 2,\ldots,n-2;$$

$$k_i = k(t_i) = \frac{vx_i ay_i - vy_i ax_i}{((vx_i)^2 + (vy_i)^2)^{3/2}} \quad i = 2,\ldots,n-2;$$

$$\varphi_i = \varphi(t_i) = k_i (vx_i^2 + vy_i^2)^{1/2} \quad i = 2,\ldots,n-2.$$

Распределения значений характеристик близки к нормальному распределению, но все же не являются таковыми и имеют более «толстые хвосты». Такие распределения в экономических исследованиях известны как распределения Парето-Леви. Закон больших чисел не применим к таким распределениям, так как выборочные средние малоинформативны и неустойчивы [3]. На рис. 1 приведена гистограмма совместного распределения первых двух характеристик траекторий курсора.

Вычисление спектров ковариационных матриц на характеристиках траекторий приводит также к фрáктальным структурам. На рис. 2 приведена гистограмма распределения первого собственного числа ковариационной матрицы. Наиболее близкий закон распределения – логнормальный. Аналогичные распределения имеют также значения остальных собственных чисел, кроме четвертого и пятого.

Известно соотношение: $R / S = (a \cdot N)^H$, где R/S – нормированный размах; N – число наблюдений; a – константа; H – показатель Херста. Оказалось, что показатель Херста, вычисленный для временного ряда каждого собственного числа ковариационной матрицы существенно меньше 0,5, что свидетельствует о фрактальной природе последовательности спектров.

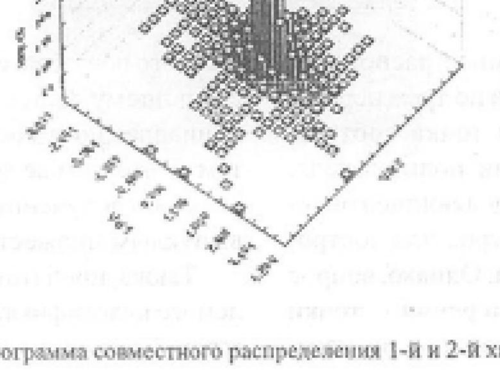

Рис. 1. Гистограмма совместного распределения 1-й и 2-й характеристик

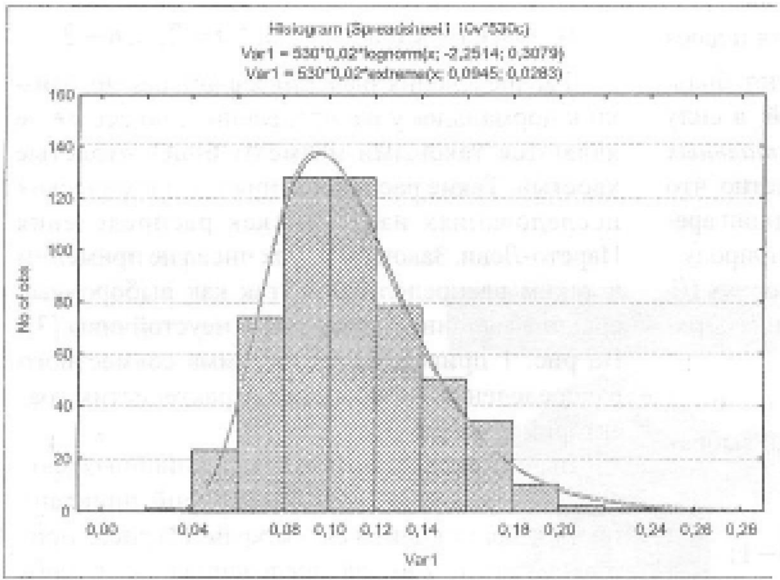

Рис. 2. Гистограмма значений первого собственного числа ковариационной матрицы

Если отобразить спектр на плоскости, где по горизонтали указан номер собственного числа, а по вертикальной оси его значение и соединить точки одного спектра отрезками рис. 3, то для построения классификатора необходимо найти метрику, которая позволит оценить расстояние между двумя спектрами в 6-ти мерном пространстве. Использование метрики $d_{12} = \max_i |\lambda_{1i} - \lambda_{2i}|$, $i = 1,...,6$, дает приемлемые результаты только для спектров всей сессии. Каждая сессия содержит более 500 сегментов – траекторий. Каждая траектория состоит в среднем из 30 точек.

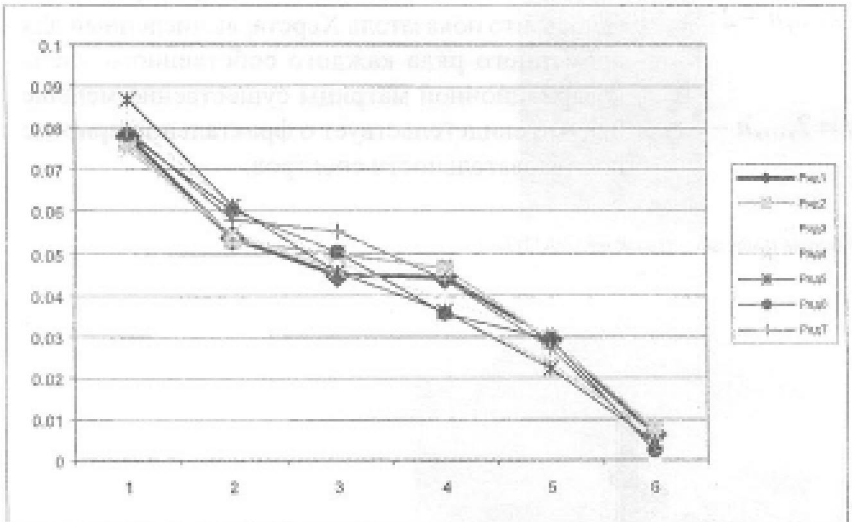

Рис. 3. Спектры ковариационной матрицы для всей сессии (жирная линия) и 6-и наиболее близких сегментов той же сессии

На рис. 4 изображено взаимное расположение спектров пяти пользователей по трем первым собственным числам. Каждая точка соответствует одной сессии измерений пользователя. Очевидно, что существуют разделяющие гиперплоскости в пространстве спектров для построения линейного классификатора. Однако, вопрос о принадлежности нового измерения – точки соответствующей испытательной сессии неизвестного пользователя, определенному уже существующему классу или же отрицание такой принадлежности остается открытым. Причиной тому – отсутствие доказательства, что множество спектров полученных для пользователя является выпуклым множеством.

Также, препятствием для построения приемлемого классификатора является нахождение некоторой величины, которая позволит нормировать

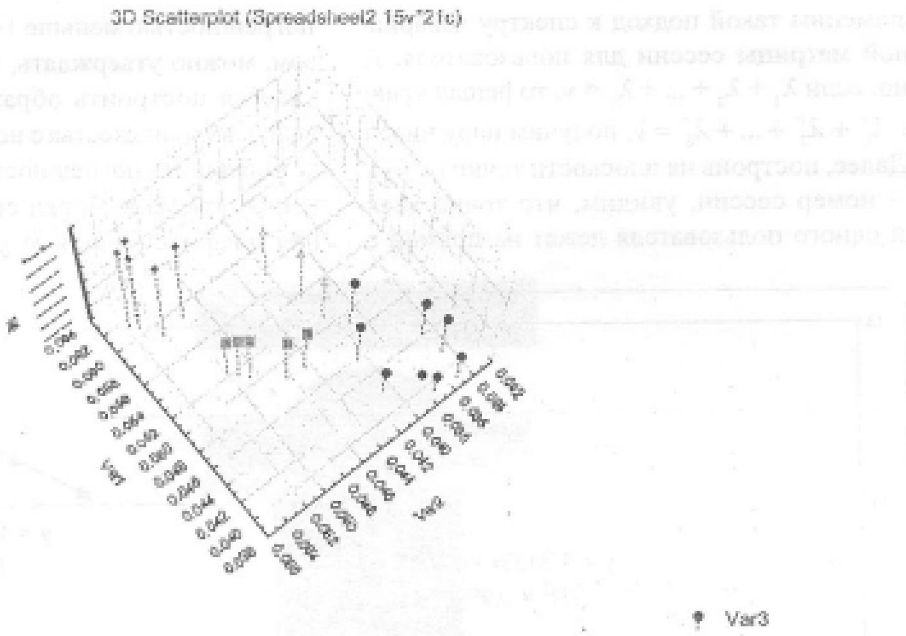

Рис. 4. Образы 5-и пользователей в трехмерном пространстве по первым трем собственным числам сессий

спектры всех пользователей. Как упоминалось выше, выборочное среднее не пригодно из-за его неустойчивости.

Чтобы попытаться решить эту проблему, рассмотрим пример. Пусть λ_1 и λ_2 некоторые положительные числа, которые связаны соотношением

$$\lambda_1 = k\lambda_2, \text{ где } k - \text{const}.$$

И пусть $\lambda_1 + \lambda_2 = \nu$. Найдем такое r, чтобы выполнялось равенство

$$\lambda_1^r + \lambda_2^r = 1.$$

Тогда

$$\nu = (1+k)(1+k^r)^{-1/r}.$$

Оказывается, что эта функция имеет ряд интересных свойств. При $0,5 < k < 1$, график функции для $0 \le r \le 2,5$ имеет вид рис. 5. Т. е., можно заменить прямой линией с точностью до четвертого знака после запятой.

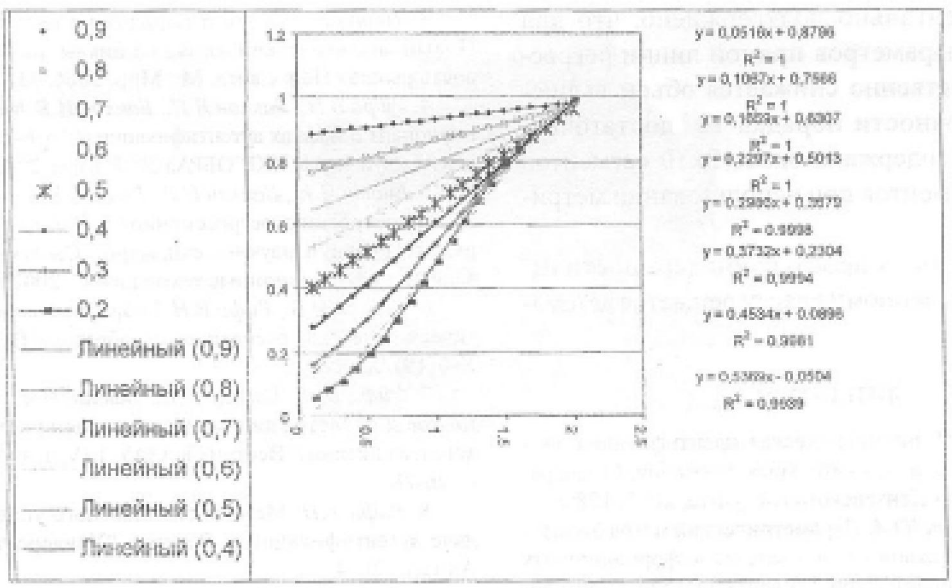

Рис. 5. Графики функции $\nu = (1+k)(1+k^r)^{-1/r}$ для различных k

Применим такой подход к спектру ковариационной матрицы сессии для пользователя. А именно: если $\lambda_1 + \lambda_2 + ... + \lambda_6 = \nu$, то решая уравнение $\lambda_1^r + \lambda_2^r + ... + \lambda_6^r = 1$, получим пару чисел (r,ν). Далее, построив на плоскости точки (r_i, ν_i), где i – номер сессии, увидим, что точки всех сессий одного пользователя лежат на прямой с погрешностью меньше $1 \cdot 10^{-3}$ рис. 6. Таким образом, можно утверждать, что в пространстве RV удается построить образ пользователя в виде выпуклого множества с некоторой погрешностью ε. Уменьшить погрешность можно за счет увеличения объема выборки сегментов для вычисления ковариационной матрицы.

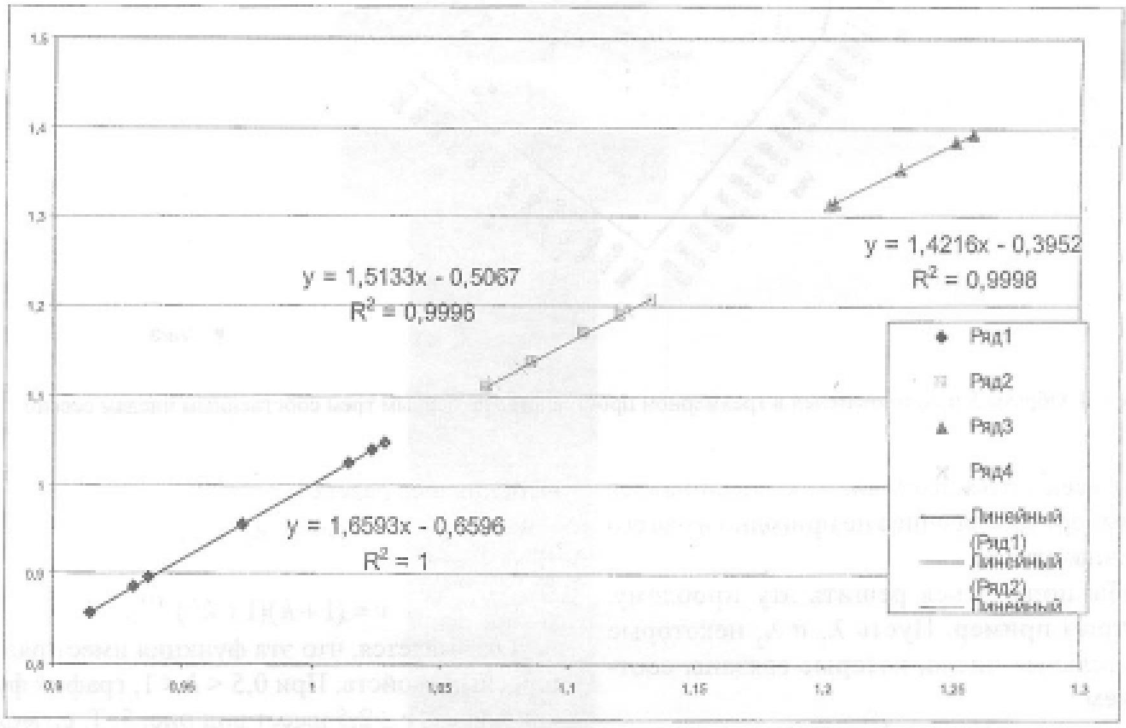

Рис. 6. Линии регрессии для трех пользователей в пространстве RV

Экспериментально подтверждено, что для определения параметров прямой линии регрессии (r,ν) существенно снижается объем вычислений. Для точности порядка 10^{-3} достаточно, чтобы сессия содержала около 3x30 сегментов вместо 500 сегментов при использовании метрики L^1.

Очевидно, что вопрос о принадлежности измерения определенному классу решается автоматически.

ЛИТЕРАТУРА

1. *Иванов А.И.* Биометрическая идентификация личности по динамике подсознательных движений. Монография. Пенза: Изд-во Пензенского гос. ун-та, 2000. 188 с.

2. *Брюхомицкий Ю.А.* Параметрический метод биометрической аутентификации пользователей информационных систем // Научно-практ. журн. "Информационное противодействие угрозам терроризма". 2003. № 1.

3. *Петерс Э.* Хаос и порядок на рынках капитала: Новый аналитический взгляд на циклы, цены и изменчивость рынка / Пер. с англ. М.: Мир, 2000. 333 с.

4. *Рифа В.Н., Баклан Я.И., Баклан И.В.* Метод главных компонент в задачах аутентификации // Тр. 6-й Всеукраинской междун. конф. УКР'ОБРАЗ 2002. Киев, 2002. С. 215-218.

5. *Бидюк П.И., Баклан И.В., Рифа В.Н.* Системный подход к построению регрессионной модели по временным рядам // Междун. научно-техн. журн. "Системные исследования и информационные технологии". 2002. № 3. К.

6. *Баклан И.В., Рифа В.Н.* Гибридные модели в статистических методах распознавания образов // Вестник ХГТУ. № 3(19). Херсон, 2003.

7. *Рифа В.Н., Баклан Я.И., Баклан И.В., Бидюк П.И., Долгов Д.С.* Метод динамических характеристик в задачах аутентификации // Вестник КАЗАУ. Т IV, ч. 4. Астана, 2004. С. 26-28.

8. *Рифа В.Н.* Методы оптимального управления в задаче аутентификации // Вестник "Университета Туран". Алматы, 2004.

9. *Рифа В.Н.* Метод динамических характеристик в задаче биометрической аутентификации. Фракталы в био-

метрии // Тез. докл. Междун. 11-й межвузовской конф. по математике и механике Евразийского нац. ун-та им. Л. Н. Гумилева. Астана, 2006.

10. *Рифа В.Н.* Метод динамических характеристик и фрактальные структуры // Междун. журн. "Системные исследования и информационные технологии". № 3. К., 2007.

11. *Рифа В.Н., Лопатин О.К.* Фрактальные структуры в задаче биометрической аутентификации // Междунар. журн. Института Проблем Искусственного Интеллекта НАНУ. № 4. Донецк, 2007р. С. 309-316.

12. *Рифа В.Н., Шарипбаев А.А.* Метод динамических характеристик и фрактального анализа в задаче динамической биометрической аутентификации // «Вестник науки» Костанайского соц.-техн. ун-та им. акад. Зулхарнай Алдамжар. 2009. № 2. С. 136-146.

13. *Рифа В.Н., Шарипбаев А.А.* Метод динамических характеристик в задаче идентификации пользователя по управлению манипулятором // Мат-лы междунар. научно-практ. конф. "Актуальные проблемы математики, информатики, механики и теории управления". Ч. 2. Алматы, 2009. С. 391-395.

Резюме

Компьютер пайдаланушының манипуляторды басқаруына байланысты ұқсастыру есебін шешуге динамикалық сипаттау әдісін қолдану мүмкіндігі қарастырылады. Сандық эксперимент нәтижесі ковариациалық матрицаның өзіндік санының бақылау саны өскен сайын пайдаланушының өзіндік тұрақтылығына жинақталатындығын көрсетеді.

Summary

The possibility of application of the dynamic characteristic method to the problem identification of a computer user by manipulator control is considered. The numerical experiments show the convergence of eigenvalues of covarianse matrixes to individual steady state values on the set of with increasing number of measurements.

Евразийский национальный университет им. Л. Н. Гумилева, г. Астана

Поступила 18.03.2010г.