Моделирование длинных биометрических кодов с зависимыми состояниями разрядов[1]

Б.С. Ахметов[1], А.И. Иванов[2], А.Ю. Малыгин[2]
[1]Казахстанский университет инновационных и телекоммуникационных систем, Уральск, Казахстан; [2]Пензенский государственный университет, Пенза, Россия
E-mail: b.akhmetov@kziitu.kz, mal890@yandex.ru

Аннотация

Рассматривается задача моделирования длинных выходных кодов нейросетевого преобразователя аналоговых (непрерывных) биометрических данных в выходной цифровой код. Дается номограмма связи регулируемых параметров моделирования с математическим ожиданием модуля коэффициентов парных корреляций.

Введение. Моделирование последовательностей независимых кодов является очень простой задачей. Так для моделирования одного кода длиной 256 бит достаточно 256 кратного обращения к соответствующему программному или аппаратному генератору случайных чисел. Далее можно попытаться найти корреляционную матрицу связи разрядов биометрического кода и построить для нее, соответствующую матрицу преобразования [1]. К сожалению, этот метод работает только для кодов низкой длины. Осуществить расчет корреляционной матрицы и соответствующего ей матричного преобразования для кодов длиной от 2 до 16 бит технически возможно. Далее задача становится некорректной. В связи с этим необходимо синтезировать матрицы связывания независимых данных по иному не классическому алгоритму.

Новая идеология связывания независимых псевдослучайных данных строится на том, что необходимо синтезировать некоторую связывающую данные матрицу так, что бы сохранялись только статистики корреляционных связей. Если точное преобразование построить невозможно, то преобразование, обеспечивающее необходимое распределение коэффициентов парных корреляций, вполне возможно.

Синтез равнокоррелированных случайных данных. Будем исходить из того, что данные некоторого вектора биометрических параметров (биометрических кодов) остаются независимыми, если связывающая их матрица является единичной:

$$\begin{bmatrix} 1 & 0 & :::::::: & 0 \\ 0 & 1 & ::::::: & 0 \\ :::: & :::: & :::::::: & ::::: \\ 0 & 0 & :::::::: & 1 \end{bmatrix} \times \begin{bmatrix} x_{1,i} \\ x_{2.i} \\ :::::: \\ x_{n,i} \end{bmatrix} = \begin{bmatrix} x_{1,i} \\ x_{2.i} \\ :::: \\ x_{n,i} \end{bmatrix} \quad \Rightarrow \quad R = \begin{bmatrix} 1 & 0 & :::::::: & 0 \\ 0 & 1 & ::::::: & 0 \\ :::: & :::: & :::::::: & ::::: \\ 0 & 0 & :::::::: & 1 \end{bmatrix} \quad (1)$$

Если же требуется создать данные с одинаковыми парными коэффициентами корреляции, то необходимо использовать связывающую матрицу с одинаковыми элементами,

[1] Статья подготовлена в рамках выполнения комплексного проекта "Разработка и подготовка производства телекоммуникационного оборудования, разработка программного сетевого, прикладного и специального обеспечения для создания цифровых сетей связи с персонализированным доступом" в соответствии с Постановлением Правительства № 218 от 09.04.2010 г.

находящимися вне единичной диагонали связывающей данные матрицы [2]:

$$\begin{bmatrix} 1 & a & :::::::: & a \\ a & 1 & ::::::: & a \\ :::: & :::: & :::::::: & ::::: \\ a & a & :::::::: & 1 \end{bmatrix} \times \begin{bmatrix} x_{1,i} \\ x_{2.i} \\ :::::: \\ x_{n,i} \end{bmatrix} = \begin{bmatrix} y_{1,i} \\ y_{2.i} \\ :::::: \\ y_{n,i} \end{bmatrix} \quad \Rightarrow \quad R = \begin{bmatrix} 1 & r & :::::::: & r \\ r & 1 & ::::::: & r \\ :::: & :::: & :::::::: & ::::: \\ r & r & :::::::: & 1 \end{bmatrix} \quad (2)$$

Элементы связывающей матрицы, находящиеся вне диагонали однозначно определяют значения парных корреляций, номограмма этой двухмерной связи $r(a, n)$ для разной длины синтезируемых векторов биометрических данных (кодов после оцифровки данных для $n = 2, 4, 8, 16, 32, 64, 128$) приведена на рис. 1.

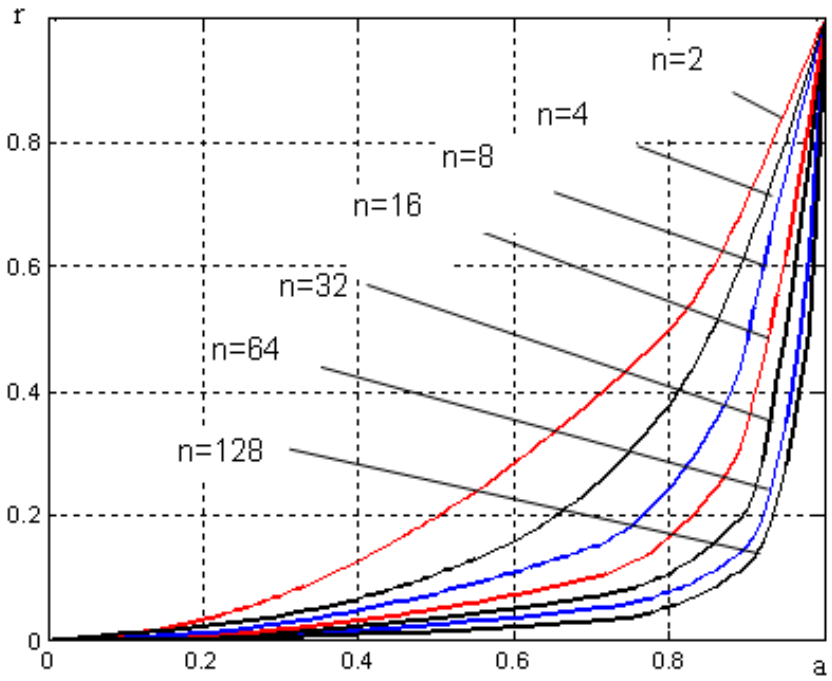

Рис. 1: Номограмма связи значений элементов связывающей матрицы «a» со значениями парных коэффициентов корреляции r моделируемых данных при длине вектора параметров $n = 2, 4, 8, 16, 32, 64, 128$

Из рисунка 1 видно, что с ростом размерности задачи значения коэффициентов парной корреляции $r(a)$ все плотнее и плотнее прижимаются к осям координат. В целом вычисление функции $r(a)$ является достаточно простой задачей, которая не требует привлечения значительных вычислительных ресурсов. Удобным является то, что изменяемый параметр «a» находится в ограниченном интервале от 0.0 до 1.0, так же как и модуль коэффициентов парной корреляции.

Моделирование данных со случайными знакопеременными коэффициентами корреляции. Практика показала, что случайная расстановка знаков $\pm$ при недиагональных элементах связывающей матрицы (2) приводит к хорошей имитации знакопеременных корреляционных связей, характерных для естественных биометрических данных. Для того чтобы дополнительно смоделировать вариации значений модулей

знакопеременных коэффициентов парной корреляции, достаточно сделать недиагональные элементы связывающей матрицы случайными [2]. Пример такого случайного распределения значений для положительных и отрицательных элементов приведен на рис. 2. В свою очередь случайное распределение значений элементов связывающей матрицы приводит к сложному распределению значений коэффициентов парной корреляции моделируемых биометрических кодов. Распределение значений парных корреляций, характерное для случайных данных рис. 2, приведено на рис. 3.

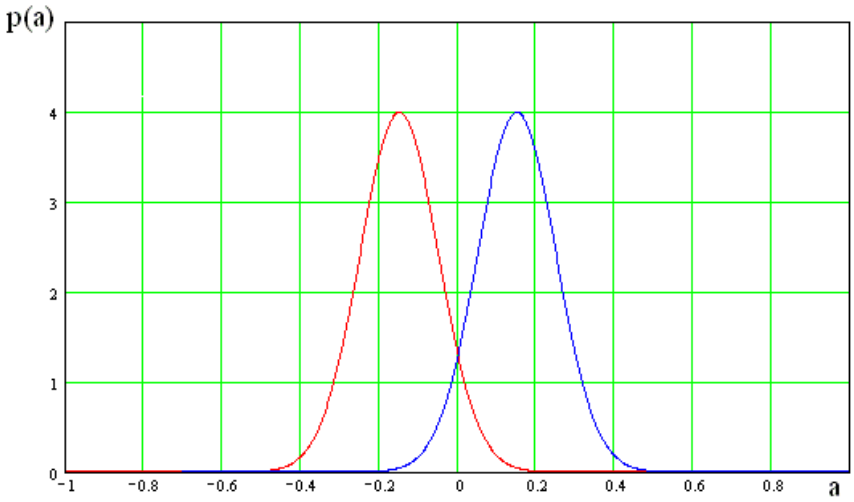

Рис. 2: Пример размывания внедиагональных элементов «$\pm a$» связывающей матрицы по нормальному закону

В конечном итоге получается, что моделирование корреляционных связей биометрических данных удобно осуществлять в обратном порядка. При этом первоначально необходимо вычислить математические ожидания модулей коэффициентов корреляции – $E(|r|)$. Далее по значению этого модуля легко определяется параметр «a» связывающей матрицы по номограмме рис. 1 или по соответствующим заранее созданным таблицам. Далее осуществляют размывание положительных и отрицательных значений внедиагональных элементов связывающей матрицы. В итоге мы получаем выходные данные с некоторой плотностью распределения значений парных коэффициентов корреляции. Пример такого распределения приведен на рисунке 3.

Регулирование среднеквадратического отклонения конечного распределения коэффициентов парной корреляции. Обычно размывание положительных и отрицательных недиагональных элементов связывающей матрицы осуществляют нормальным законом (см. рис. 2), однако для этой цели может быть использован и любой иной генератор случайных чисел. Размывание следует осуществлять, постепенно увеличивая среднеквадратическое отклонение нормальных законов распределения с заранее заданными математическими ожиданиями $\pm a$. При этом следует контролировать размывание коэффициентов парной корреляции моделируемых биометрических данных. Существует однозначная монотонная связь между дисперсиями двух законов распределения значений параметров связывающей матрицы «a» и «$-a$» со среднеквадратическим

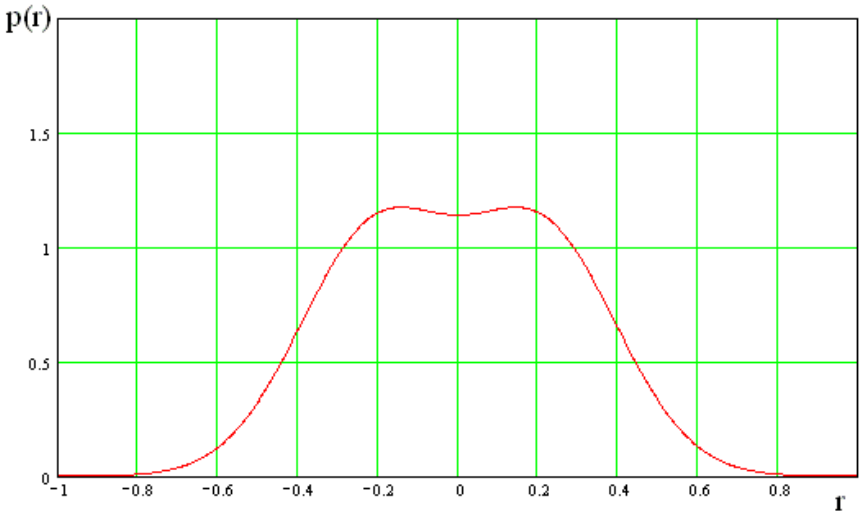

Рис. 3: Распределение значений парных коэффициентов корреляции, полученное при использовании данных двух симметричных генераторов параметра «a» связывающей матрицы

отклонение модулей коэффициентов парной корреляции $\sigma\left(|r|\right)$.

Если положительные и отрицательные параметры связывания «a» имеют одинаковые (симметричные относительно точки $r = 0.0$) распределения, то и плотность распределения $p(r)$, будет являться симметричной относительно точки $r = 0.0$. В этом случае среднеквадратическое отклонение $\sigma(r)$ распределения $p(r)$ является трехмерной функцией:

$$\sigma(r) = f(a, \sigma(a), n). \tag{3}$$

Частные производные трехмерной функции (3) по каждому из ее параметров являются аналитическими и имеют один знак, что значительно облегчает регулирование среднеквадратического отклонения распределения коэффициентов парной корреляции. В конечном итоге моделирование сводится к последовательному обращению к нескольким номограммам подобным номограмме рис. 1 и многократному запуску n генераторов псевдослучайных независимых данных с нормальным законом распределения значений.

Заключение. Описанный выше метод моделирования позволяет надежно воспроизводить усредненные характеристики биометрических данных и биометрических кодов, если эти данные оцифровать. С помощью этого метода нельзя добиться заранее заданной корреляционной матрицы моделируемых данных, однако по распределениям контролируемых статистических параметров (например, по распределениям значений коэффициентов корреляций или модулей коэффициентов корреляций) данные хорошо совпадают. В конечном итоге мы получаем распределение генератора биометрических данных (биометрических кодов) с возможностью регулировать математическое ожидание и среднеквадратическое отклонение распределения значений коэффициентов парной корреляции биометрических данных и биометрических кодов.

Список литературы

[1] *Шалыгин А.С., Палагин Ю.И.* Прикладные методы статистического моделирования. – Л.: Машиностроение, 1986 г. – 320 с.

[2] *Малыгин А.Ю., Волчихин В.И., Иванов А.И., Фунтиков В.А.* Быстрые алгоритмы тестирования нейросетевых механизмов биометрико-криптографической защиты информации. – Пенза: Изд-во Пензенского ГУ, 2006 г. – 161 с.

Akhmetov B.S., Ivanov A.I., Malygin A.Yu., Biometric image converter biometrics-code dependent data modeling, The Bulletin of KazNU, ser. math., mech., inf. 2011, №3(70), 57 – 61

Long weekend is neural network modeling codes converter analog (continuous) biometric data into the output code. A nomogram of controlled parameters simulation with mathematical expectation pairwise correlation coefficients module.

Б.С. Ахметов, А.И. Иванов, А.Ю. Малыгин, Дәреженелік жағдайына байланысты ұзын биометриялық кодтарды үлгілеу, ҚазҰУ хабаршысы, мат., мех., инф. сериясы 2011, №3(70), 57 – 61

Ұзын қосалқы нейрожелілік үйлесімді (үздіксіз) биометриялық негізгі сандық кодтың есебін үлгілеу қаралады. Математикалық коэффициент модулін күту екі корреляциялық реттегіш параметрлеріне байланысты номограмма беріледі.