

УДК 517(075)

С. В. Белинский (6 курс, каф. САиУ), Б. И. Морозов, к.т.н., доц.

ИДЕНТИФИКАЦИЯ ЛИЧНОСТИ ПО КЛАВИАТУРНОМУ ПОЧЕРКУ

Развитие различных видов коммуникаций на базе персональных компьютеров привело во многих случаях к необходимости идентификации личности. Наряду с широко распространенными методами идентификации (исследование отпечатков пальцев, строение радужной оболочки глаза, идентификация голоса, почерковедческая экспертиза текста и подписи) во многих случаях доступным для исследования является лишь клавиатурный почерк.

Объективно характеристиками клавиатурного почерка являются продолжительность удержания определенной клавиши, сила нажатия, длительность паузы между последовательными нажатиями. Предполагается, что указанные параметры в статистическом смысле являются индивидуальными. Использование для идентификации усилия нажатия на клавишу требует специального датчика, не предусмотренного на стандартной клавиатуре. Кроме того, физически обоснованно предполагать, что данное усилие связано с длительностью нажатия. В работе делается попытка идентификации личности по двум оставшимся параметрам, что не требует специальной аппаратной доработки, а вместо этого ставит перед нами задачу программной реализации идеи.

Представляется, что индивидуальные биометрические особенности (осознание передаваемого слова, сравнения с правилами правописания, реакция, координация мышечно-двигательного аппарата пальцев, выработанная манера печати, особенности расположения клавиш на клавиатуре) являются индивидуальными и пригодными для статистической обработки в целях решения задачи. Поэтому в качестве параметров являющихся функциями выше перечисленных биометрических особенностей выбираются длительность нажатия на клавиши и длительность паузы между нажатиями каждой возможной пары букв текста.

В результате накопления статистических данных вычисляются математические ожидания и среднеквадратические отклонения по каждому параметру для каждого клиента. Задача определения клиента по данной (по текущей) выборке принятых параметров формулируется в терминах теории правдоподобия гипотез. Под гипотезой S_j понимается принадлежность полученной выборки j -ому клиенту. Решение принимается по максимуму апостериорной вероятности гипотезы S_j при условии получения выборки Z_1, Z_2, \dots, Z_n по формуле Байеса.

$$P(S_j | Z_1, Z_2, \dots, Z_n) = \frac{p(S_j) * W(Z_1, Z_2, \dots, Z_n | S_j)}{\sum_{i=1}^m p(S_i) * W(Z_1, Z_2, \dots, Z_n | S_i)}, \quad (1)$$

где: $p(S_k)$ - априорная вероятность гипотезы S_k ; $W(Z_1, Z_2, \dots, Z_n | S_k)$ - совместная плотность распределения параметров вектора Z для k -ого клиента; m - кол-во клиентов, n - количество параметров.

Решение принимается в пользу того j , для которого выражение в формуле (1) имеет максимальное значение. Значение вероятности по формуле (1) есть вероятность правильности идентификации клиента.

В работе показано, что при нормальном распределении параметров Z_1, Z_2, \dots, Z_n , при условии их независимости, априорно равно вероятности гипотез S_1, S_2, \dots, S_m , а также в

предположении равенства среднеквадратических отклонений параметров Z_1, Z_2, \dots, Z_n критерий Байеса совпадает с критерием Махалобиса-Хемминга - нахождения минимума выражения E по формуле:

$$E = \sqrt{\sum_{j=1}^k (v_j - m(v_j))^2}$$

Кроме этого сформулированы требования к периоду заполнения тактовыми импульсами счетчика и показано, что вероятность определения правдоподобия гипотезы по формуле (1) может быть сделана сколь угодно близкой к единице вне зависимости от количества клиентов за счет увеличения длины выборки. На практике при получении коротких сообщений длина выборки может оказаться недостаточной для получения достоверного решения идентификации личности (появление конкурирующих гипотез). Указанная ситуация является неизбежной платой за попытку принятия решения в условиях ограниченной информации. Для эффективного использования данной программы необходима программная реализация ее работы в поле часто используемых программ операционной системы.

Данная программа реализована с использованием Assembler, Delphi и работает под управлением ОС WIN98.