
СЗИ α_i с помощью полученной модели рассчитывается фактическая защищенность ИВС $P_{\text{дн}}(\alpha_i)$. Показатель уязвимости ИВС рассчитывается по формуле

$$Q_{\text{дн}}(\alpha_i) = 1 - P_{\text{дн}}(\alpha_i).$$

Таким образом, рассчитав показатель уязвимости ИВС, используя метод оценки уязвимости ИВС по степени ее оснащённости СЗИ, можно сделать вывод о фактической защищённости ИВС в процессе выполнения мероприятий по защите ИВС. Это позволяет делать выводы об эффективности внедрения или модификации подсистемы защиты. Безусловно, разработанный метод не отменяет необходимость использования инструментальных средств анализа уязвимостей, конкретных программных и/или аппаратных средств, а только дополняет их, в качестве обобщенного показателя уязвимости ИВС.

СПИСОК ЛИТЕРАТУРЫ:

1. Боридько С. И., Тихонов Б. Н. Модели управления состоянием критически важных объектов: монография. М.: МИНИТ ФСБ России, 2006.
2. Домарев В. В. Безопасность информационных технологий. Системный подход. К.: ООО «ТИД Диа Софт», 2004.

А. Н. Голубинский

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ РЕЧЕВОГО СИГНАЛА, ОСНОВАННАЯ НА АППРОКСИМАЦИИ СПЕКТРА НАБОРОМ ПОСТОЯННЫХ СОСТАВЛЯЮЩИХ В СООТВЕТСТВУЮЩИХ ПОЛОСАХ ЧАСТОТ

Введение

В современных системах безопасности и информационных системах можно выделить две важные научно-практические задачи — идентификация и верификация личности [1]. При верификации (подтверждении) личности человека требуется установить его соответствие данному эталону, приняв одно из двух решений: заявитель является тем, за кого он себя выдает, или не является. При идентификации (установлении) личности человека необходимо выбрать из имеющейся базы данных эталонов тот эталон, на который заявитель максимально похож, при этом нужно принять решение: заявитель наиболее похож на конкретную персону (чей эталон находится в базе данных) или заявитель не соответствует ни одной из персон (имеющихся в базе данных).

В последнее время все более часто находят применение биометрические системы аутентификации (верификации и идентификации) личности [1], принцип работы которых основывается на анализе различных персональных физиологических характеристик людей, таких как форма и размеры руки, отпечаток пальца, голос, параметры зрачка и сетчатки глаза, форма и размеры лица и т. д. Одним из перспективных способов аутентификации личности является подтверждение или установление личности по голосу на основе речевого сигнала человека [1].

Следует отметить, что существуют различные методы построения моделей речевых сигналов [2, 3]. Приведем пять основных подходов к созданию математических моделей речевых сигналов, заданных функциональной зависимостью отсчетных значений модели от времени:



1) линейное предсказание речевого сигнала; недостатками данного подхода являются: малое время предсказания, а также тот факт, что получение отсчетных временных значений сигнала математической модели основано на использовании дополнительного источника сигнала в виде белого шума, что влечет за собой определенные трудности [2, 4];

2) решение дифференциальных уравнений (модифицированное волновое уравнение и граничные условия) [3], составленных при использовании нескольких условных физических приближений, эти уравнения описывают распространение акустического колебания в речевом тракте, с учетом возбуждающего сигнала; в данном подходе присутствуют следующие основные недостатки: при аппроксимации возбуждающего сигнала и площади поперечного сечения речевого аппарата если и удастся, как правило, численными методами решить данное волновое уравнение, то полученные результаты имеют очень сложный характер пространственно-временной зависимости и малоприспособлены для быстрого и адекватного анализа речи;

3) гармоническая математическая модель речевого сигнала, использующая для представления сигнала синусоиды кратной частоты (оставшаяся часть сигнала, которая не может быть описана при помощи данной модели, называется шумовой). Представление сигнала в форме гармоник плюс шум эффективно используется во многих речевых приложениях [5, 6]. Следует отметить, что от разделения математической модели речевого сигнала на периодическую (вокализованные участки речи) и шумовую (невокализованная речь) части в существенной степени зависит адекватность модели и ее применимость в конкретных технических задачах [2]. Данную модель можно синтезировать, используя, например, коэффициенты дискретного преобразования Фурье речевого сигнала (также можно применять другие формы спектральных или кепстральных коэффициентов). Основные недостатки: использование, как правило, большого числа коэффициентов модели, а также ограниченная точность модели, связанная с бесконечной длительностью гармонической функции;

4) математическая модель речевого сигнала в виде набора отрезков функций, синтезируемая на основе вейвлет-коэффициентов [7];

5) построение математической модели речевого сигнала в виде явной функциональной временной зависимости, основанное на теории модуляции (как частный случай, модулированный импульс [8, 9]).

Отметим, что точность верификации и идентификации личности по голосу будет тем выше, чем точнее модель, описывающая речевой сигнал. Очевидные требования, предъявляемые к выбору модели речевого сигнала при одинаковой точности модели, — наименьшее количество коэффициентов модели и ее простота, выполнение данных требований способствует уменьшению систематической ошибки и времени обработки поступивших данных.

Таким образом, представляет научный интерес разработка моделей речевого сигнала, отражающих индивидуальные особенности голоса человека, его уникальность при произнесении речевого сообщения.

Известно, что наиболее эффективно проведение верификации и идентификации личности по голосу на основе речевого материала, содержащего вокализованные участки речи, а именно гласные звуки [10].

Отметим, что одной из самых простых форм описания речевых сигналов является представление спектра в виде набора постоянных составляющих в соответствующих полосах частот, которыми обычно являются формантные частоты [11]. Данный способ описания речевых сигналов успешно применяется для задач идентификации и верификации диктора по голосу [11].

Исследуем возможность построения математической модели речевого сигнала на вокализованные участки речи по описанному выше способу, при этом будем основываться на аппроксимации спектра набором постоянных составляющих в соответствующих полосах частот. В виде данных

полос частот будем использовать окрестности относительно частоты основного тона и обертонов речевого сигнала (частоты гармоник импульсного периодического сигнала, генерируемого голосовыми связками) в области первых двух [12] формантных частот (частоты глобальных резонансов речевого тракта).

Цель работы — разработка методики построения и расчета параметров математической модели речевого сигнала, основанной на аппроксимации спектра набором постоянных составляющих в соответствующих полосах частот, пригодных для аутентификации личности по голосу.

1. Теоретический анализ

Положим, что речевой материал, используемый для аутентификации личности по голосу, содержит вокализованные участки речи. Поэтому для построения математической модели речевого сигнала в виде некоторой зависимости от времени будем использовать подход, основанный на выделении модулирующей (информационной) и модулируемой (несущей) компонент [3, 9].

Воспользуемся детерминированным подходом к построению математической модели (речевого сигнала, содержащего вокализованные участки речи), основанной на теории модуляции. Аппроксимируем спектральную плотность $S(\omega)$ [В·с] речевого сигнала набором постоянных составляющих S_l , $l = \overline{0; L}$ в полосах частот шириной Π_l [Гц] в окрестности несущих частот $f_l^{\text{нес}}$ [Гц] (постоянной составляющей частоты, частоты основного тона f_0 и частот обертонов f_1, f_2, \dots, f_{L-1}) речевого сигнала.

$$S(2\pi f) \equiv S(\omega) = \begin{cases} S_0, \text{ при } f \in [0; f_0^{\text{нес}} + \frac{\Pi_0}{2}) \\ S_1, \text{ при } f \in [f_1^{\text{нес}} - \frac{\Pi_1}{2}; f_1^{\text{нес}} + \frac{\Pi_1}{2}) \\ \dots \\ S_L, \text{ при } f \in [f_L^{\text{нес}} - \frac{\Pi_L}{2}; f_L^{\text{нес}} + \frac{\Pi_L}{2}] \end{cases} \quad (1)$$

Математическую модель речевого сигнала в виде детерминированной функции (в явном виде зависящей от времени) получим, вычислив обратное преобразование Фурье от спектральной плотности (1). В результате математическую модель можно привести к виду:

$$u(t) = \sum_{l=0}^L U_l \text{sinc}(2\pi F_l t) \cos(2\pi f_l^{\text{нес}} t), \quad (2)$$

где $F_l = \Pi_l/2$; $U_l = S_l \Pi_l = 2S_l F_l$ [В]; $\text{sinc}(x) = \sin(x)/x$; $[u(t)] = \text{В}$.

Следует отметить важный частный случай модели (2) при выполнении условий:

1) $f_l^{\text{нес}} = l f_0$, $l = \overline{0; L}$ (частоты обертонов кратны частоте основного тона);

2) $\Pi_l = \Pi_0$ (полосы частот равны между собой по ширине);

при данных допущениях математическая модель упрощается, преобразуясь к виду:

$$u(t) = \text{sinc}(2\pi F_0 t) \sum_{l=0}^L U_l \cos(2\pi l f_0 t). \quad (3)$$

Из соотношения (3) видно, что данная математическая модель речевого сигнала является амплитудно-модулированным колебанием:

$$u(t) = u_{\text{мод}}(t) \cdot u_{\text{нес}}(t), \quad (4)$$

в котором можно условно выделить: $u_{\text{мод}}(t) = \text{sinc}(2\pi F_0 t)$ [безразмерная величина] — модулирующее колебание и $u_{\text{нес}}(t) = \sum_{l=0}^L U_k \cos(2\pi l f_0 t)$ [В] — несущее колебание.



2. Методика

Рассчитаем коэффициенты предложенной математической модели (3) речевого сигнала. Для расчета коэффициентов U_l , $l = \overline{0, L}$, f_0 , F_0 , входящих в модель (3), будем использовать персональную ЭВМ, ввод речевого сигнала в которую выполняется с помощью звуковой платы со стандартной частотой дискретизации $f_d = 6000$ Гц. Данная частота была выбрана вследствие того, что первые 3–4 форманты находятся в области до 3000–3600 Гц [2]. Звуковые платы ЭВМ, как правило, добавляют различные постоянные составляющие в речевой сигнал.

В связи с этим проводилась операция центрирования реализации: $y_i = x_i - \bar{x}$, $i = \overline{1, N}$, где x_i [В] – начальные отсчеты речевого сигнала; $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$ – математическое ожидание; N – число отсчетов.

Длительность импульса (время произнесения парольной фразы) была задана $\tau_{\text{и}} = 0,3$ с, при данном интервале дискретизации $\Delta = 1/f_d$ речевой сигнал имеет $N = 1800$ отсчетов.

Для определения коэффициентов модели (3) в качестве степени несоответствия модели и оригинала была принята ошибка модели между коэффициентами корреляции (КК) R_j [безразмерная величина] центрированного речевого сигнала и значениями нормированной автокорреляционной функции (АКФ) Ra_j [безразмерная величина] математической модели (3):

$$\varepsilon R = \sum_{j=1}^J (R_j - Ra_j)^2, \quad (5)$$

где $R_j = K_j / K_0$; здесь функция корреляции (ФК) речевого сигнала:

$$K_j = \frac{1}{N} \sum_{i=1}^{N-j} (y_i - \bar{y})(y_{i+j} - \bar{y}) \text{ [B}^2\text{]}; \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i, \quad j = \overline{0, J};$$

J – число отсчетов КК (для центрированной реализации $\bar{y} = 0$); нормированная АКФ $Ra_j = Ra(j\Delta)$:

$$Ra(\tau) = \frac{Ka(\tau)}{Ka(0)} = \frac{Ka(\tau)}{Ea}, \quad (6)$$

где $Ka(\tau)$ [B²·с] – АКФ модели речевого сигнала, которая рассчитывается с помощью обратного преобразования Фурье от спектральной плотности энергии $W(\omega) = |S(\omega)|^2$ [B²·с²]. Для модели (3) речевого сигнала получим для $\tau \geq 0$:

$$Ka(\tau) = \frac{\text{sinc}(2\pi F_0 \tau)}{2F_0} \sum_{l=0}^L (U_l)^2 \cos(2\pi l f_0 \tau); \quad (7)$$

$Ea = \frac{1}{2F_0} \sum_{l=0}^L (U_l)^2$ [B²·с] – энергия математической модели речевого сигнала.

3. Экспериментальная часть

Определим число отсчетов коэффициентов корреляции J , достаточное для получения адекватных характеристик модели. Анализ речевых сигналов различных дикторов показал, что все значения коэффициентов корреляции заходят в доверительные границы Бартлетта $\pm 3\sigma$ до 200 отсчета для частоты дискретизации $f_d = 6000$ Гц, таким образом, положим, что $J = 200$. Здесь $\sigma = \sqrt{\frac{1}{N} \left(1 + 2 \sum_{k=1}^K (R_k)^2 \right)}$ [В] – стандартная ошибка для коэффициентов корреляции; N – количество отсчетов речевого сигнала; K – некоторое значение количества отсчетов коэффициентов корреляции, не превышающее примерно $N/4$.

Частоту основного тона будем оценивать при помощи методики, основанной на определении минимума невязки коэффициентов корреляции [9]. Невязка определяется между значениями коэффициентов корреляции, полученными на основе экспериментальных данных, и коэффициентов тестовой (специальной для оценки основного тона) математической модели речи, содержащей



вокализованные участки. Оценка \hat{f}_0 частоты основного тона f_0 определяется как значение аргумента, при котором наблюдается наименьшее значение невязки $\epsilon RTest(f_0)$ в диапазоне частот f_0 : от 50 до 350 Гц (с шагом, например, $\Delta f_0 = 0,1$ Гц):

$$\hat{f}_0 = \arg \inf [\epsilon RTest(f_0)], \quad (9)$$

здесь тестовая невязка (ошибка):

$$\epsilon RTest(f_0) = \frac{1}{J} \sum_{j=1}^J (RaTest(j\Delta, f_0) - R_j)^2, \quad (10)$$

где $RaTest(j\Delta, f_0)$ — коэффициент корреляции тестовой математической модели речевого сигнала, применяемый для оценки частоты основного тона, упрощенный вид которого:

$$RaTest(j\Delta, f_0) = \frac{1}{3} \sum_{l=0}^3 \cos(2\pi l f_0 j \Delta). \quad (11)$$

С использованием данной оригинальной методики становится возможным без применения сложных алгоритмов вычисления получить высокоточную оценку частоты основного тона \hat{f}_0 речевого сигнала (содержащего вокализованные участки речи).

Рассчитаем коэффициенты модели U_l , для этого согласно методу наименьших квадратов минимизируем ошибку модели ϵR (5), приравняв к нулю первые производные ϵR по каждому из коэффициентов, в результате получим следующую систему, состоящую из $(L+1)$ уравнений:

$$\begin{cases} \sum_{j=1}^J (R_j - Ra_j) \frac{\partial Ra_j}{\partial U_0} = 0; \\ \sum_{j=1}^J (R_j - Ra_j) \frac{\partial Ra_j}{\partial U_1} = 0; \\ \dots \\ \sum_{j=1}^J (R_j - Ra_j) \frac{\partial Ra_j}{\partial U_L} = 0. \end{cases} \quad (12)$$

Уравнения данной системы представляют собой полиномы 4-й степени, вследствие чего система однозначного решения не имеет. Положим, что энергия одинакового речевого материала — величина постоянная ($Ea = const$), при использовании этого допущения система нелинейных уравнений (12), состоящая из полиномов четвертой степени, сводится к линейной (из-за постоянного знаменателя в выражении (6) для Ra (τ) и имеет однозначное решение (при введении удобной замены переменных $V_l = U_l^2$ [В²]).

4. Результаты

В качестве примера применения разработанной математической модели (3) для задач аутентификации по голосу будем использовать речевой материал в виде слова «он», состоящего из одного гласного звука [о] и одного сонорного согласного звука [н]. При заданных параметрах: $f_d = 6000$ Гц; $L = 5$; $F_0 = 10$ Гц ($\Pi_0 = 20$ Гц); $Ea = 1$ В²·с; $J = 200$ — были получены следующие семь коэффициентов модели, отражающих индивидуальные особенности голоса: $f_0 = 155,2$ Гц; $U_0 = 0,929$ В; $U_1 = 1,696$ В; $U_2 = 1,629$ В; $U_3 = 3,022$ В; $U_4 = 0,664$ В; $U_5 = 0,681$ В. При этом ошибка модели составила $\epsilon R = 1,086$. Численный сравнительный анализ показывает, что нормированная АКФ модели сигнала (3) достаточно точно аппроксимирует КК речевого сигнала. Данное обстоятельство указывает на пригодность предложенной модели к описанию гласных и сонорных согласных звуков.



При использовании предложенной модели речевого сигнала (3) необходимо вычисление и хранение в банке данных семи коэффициентов ($f_0, U_0, U_1, U_2, U_3, U_4, U_5$), входящих в модель сигнала (3), для каждого из верифицируемых дикторов.

Коэффициенты разработанной модели (для $\Pi_0 = 20$ Гц), с точностью до постоянного множителя, практически совпадают с коэффициентами модели в виде импульса АМ колебания с несколькими несущими частотами [7] (при одинаковых энергиях сигналов двух моделей). Таким образом, можно проследить довольно точную взаимосвязь между коэффициентами математических моделей речевого сигнала, в одной из которых содержится параметр речевого сигнала в виде конечной длительности $\tau_{\text{и}}$ [7], а в другой (3) этот параметр отсутствует, так, модулирующее колебание представляет собой быстро затухающую осциллирующую функцию вида $\sin(x)/x$. Относительные рассогласования $\delta U_l = \left(\frac{|U_l^{\text{II}} - U_l^{\text{I}}|}{U_l^{\text{I}}} \right) \cdot 100\%$ между коэффициентами первой и второй моделей соответственно равны: $\delta U_0 = 39,093\%$; $\delta U_1 = 0,025\%$; $\delta U_2 = 0,177\%$; $\delta U_3 = 0,041\%$; $\delta U_4 = 0,337\%$; $\delta U_5 = 0,508\%$. Как видно из результатов анализа, для коэффициентов $l \in \overline{1;5}$ рассогласование не превышает 1%. Рассогласование между амплитудами низкочастотной составляющей (нулевой коэффициент) двух моделей в виде 39% обусловлено структурой второй модели (без учета длительности). Однако в задачах аутентификации личности по голосу вклад данного нулевого коэффициента сравнительно мал относительно остальных пяти амплитудных коэффициентов, и данную модель (3) можно использовать для задач идентификации и верификации, особенно (ввиду простой формы модели) на этапах предварительного анализа.

Следует отметить, что предложенная математическая модель речевого сигнала адекватно и с высокой степенью точности описывает вокализованные участки речи, при этом модель является довольно простой в анализе и расчете ее существенных параметров, четко прослеживается ее физический смысл. Предложенную математическую модель речевого сигнала (3) целесообразно применять в тех случаях, когда затруднена оценка длительности произнесения контрольного речевого материала, т. е. затухающая огибающая модулирующего (информационного) колебания дает возможность без оценки длительности речевого сигнала осуществлять адекватное описание вокализованных участков речи.

Таким образом, разработана методика расчета параметров математических моделей речевого сигнала, основанных на аппроксимации спектра набором постоянных составляющих в соответствующих полосах частот. Предложенная модель речевого сигнала (3) для аутентификации личности по голосу адекватно описывает речевые сигналы, содержащие вокализованные участки речи.

СПИСОК ЛИТЕРАТУРЫ:

1. Сорокин В. Н. Фундаментальные исследования речи и прикладные задачи речевых технологий // Речевые технологии. 2008. № 1. С. 18–48.
2. Назаров М. В., Прохоров Ю. Н. Методы цифровой обработки и передачи речевых сигналов. М.: Радио и связь, 1985. — 176 с.
3. Сорокин В. Н. Синтез речи. М.: Связь, 1992. — 392 с.
4. Ролдугин С. В., Голубинский А. Н., Вольская Т. А. Модели речевых сигналов для идентификации личности по голосу // Радиотехника. 2002. № 11. С. 79–81.
5. Stylianou Y. Apply the harmonic plus noise model in concatenative speech synthesis // IEEE Trans. on Speech and Audio Process. 2001. Vol. 9. № 1. P. 21–29.
6. Zavarehei E., Vaseghi S., Yan Q. Noisy speech enhancement using harmonic-noise model and codebook-based post-processing // IEEE Trans. on Speech and Audio Process. 2007. Vol. 15. № 4. P. 1194–1203.
7. Рассказова С. И., Власов А. И. Метод формантного анализа на основе вейвлет-преобразования в системах распознавания речи // IX Научно-техническая конференция «Наукоемкие технологии и интеллектуальные системы»: Сборник трудов. Москва: МГТУ им. Н. Э. Баумана, 2007. С. 38–43.
8. Якушев Д. И., Скларов О. П. Моделирование гласных звуков // Акустический журнал. 2003. Т. 49. № 4. С. 567–569.
9. Голубинский А. Н. Методика расчета параметров модели речевого сигнала в виде импульса АМ-колебания с несколькими несущими частотами, для случая модуляции суммой гармоник // Системы управления и информационные технологии. 2008. № 4.1. С. 156–161.



10. Аграновский А. В., Леднов Д. А., Репалов С. А. Метод текстонезависимой идентификации диктора на основе индивидуальности произношения гласных звуков // Акустика и прикладная лингвистика: Ежегодник РАО. 2002. Вып. 3. С. 103–115.

11. Патент РФ № 2230375: МПК G 10 L 15/00, G 10 L 17/00. Метод распознавания диктора и устройство для его осуществления / П. В. Лабутин, А. Н. Раев, С. Л. Коваль – № 2002123509/09; заявл. 03.09.02; опубл. 10.06.04.

12. Чистович Л. А., Венцов А. В., Грамстрем М. П. и др. Физиология речи. Восприятие речи человеком. М.: Наука, 1976. – 388 с.

А. А. Дураковский

УПРАВЛЕНИЕ РАСПРЕДЕЛЕННОЙ СИСТЕМОЙ ЗАЩИТЫ ДОСТУПА К ДАННЫМ

Распространение вычислительных сетей предоставляет возможности для несанкционированного доступа к информационным системам (ИС), а тенденция к переходу на распределенные вычислительные системы уменьшает возможности централизованной защиты доступа к данным.

Прежде чем приступать к построению системы управления распределенной системой защиты доступа к данным, следует определить границы системы информационной безопасности [1–6]. Это можно сделать на основе следующих данных:

- структура организации — описание существующей структуры и тех изменений, которые предполагается внести в связи с разработкой системы информационной безопасности;
- размещение методов и средств вычислительной техники и поддерживающей инфраструктуры;
- ресурсы информационной системы, подлежащие защите. Рекомендуется рассмотреть ресурсы автоматизированной системы, данные, системное и прикладное программное обеспечение. Поскольку для организации все эти ресурсы представляют ценность, должна быть выбрана система критериев и методология получения оценок по этим критериям;
- технология обработки информации и решаемые задачи. Для решаемых задач следует построить модели обработки информации в терминах ресурсов.

В результате определяются технические требования к системе информационной безопасности, в которых фиксируются границы системы, перечисляются подлежащие защите ресурсы и дается система критериев для определения их ценности.

На комплексную систему технической защиты информации нормативными документами возлагается задача обеспечения функциональных свойств защищенных информационных систем. Эта задача решается как техническими, так и программными средствами базового и прикладного программного обеспечения, а также с использованием специально разрабатываемых программных и аппаратных средств технической защиты информации.

Для реализации организационно-правовых мероприятий в большинстве случаев нет необходимости использования средств, являющихся компонентами информационных систем.

Основная задача технических мероприятий — обеспечение как физической, так и информационной безопасности.

Информационная безопасность обеспечивается использованием технических средств для [7, 10]:

- построения модели защищенной системы;
- управления доступом к ресурсам информационных систем;
- обеспечения целостности и конфиденциальности;

