УДК 621.391.176

ПОВЫШЕНИЕ ПОМЕХОУСТОЙЧИВОСТИ ПРИЕМА В СИСТЕМАХ ПЕРЕДАЧИ ДАННЫХ

к.т.н. О.П. Малофей, А.О. Малофей*
(представил проф. А.В. Королев)

Предложен подход к повышению помехоустойчивости приема при помощи процедуры обучения с выровненными оценками в системах передачи данных.

Известно, что размер зоны неопределенности принятия решения первой решающей схемы (РС) влияет на помехоустойчивость приема [1], так как уровень порога стирающего канала и размер зоны стирания однозначно связаны с вероятностями трансформации и ложного стирания.

Метод повышения помехоустойчивости приема в условиях неопределенности, предложенный Чейзом [1], состоит в упорядочении принятых символов по их надежности и выборе из их числа 5 - 10 наименее надежных символов для специальной обработки. Такая процедура названа "чейзингом" и применяется, в основном, для декодирования кодов Боуза - Чоудхури - Хоквингема (БЧХ - кодов) в каналах связи (КС) с достаточно слабым уровнем шумов. Чейзинг представляет собой общий метод, с помощью которого характеристики почти любого алгоритма декодирования блоковых кодов можно улучшить за счет дополнительных затрат времени на вычисления. При этом увеличение на 2 – 3 количества подвергаемых чейзингу символов вызывает необходимость в 4 – 8 раз ускорять работу декодера, чтобы конечное время доведения информации осталось неизменным.

Однако данный метод принятия нежестких решений не решает проблему выбора и регулирования уровня порога в процессе приема информации в реальных КС АСУ, подверженных воздействию мощных слабо коррелированных помех, ввиду вероятности появления чрезмерно большого количества стертых символов, либо замены запрещенных комбинаций разрешенными в процессе чейзинга. При этом ошибки конечного результата декодирования обусловлены неоптимальностью выбора порога в первой РС.

Некоторые методы выбора уровня порога РС ДСК рассмотрены в [2] и сводятся к решению задачи оптимизации системы. К их числу, как наиболее применимые в настоящее время, следует отнести методы оптимизации уровня порога по критериям Неймана - Пирсона, а также по минимуму и максимуму [3]. Однако, практическое применение данных методов в реальных КС не позволяет достичь рассчитанного теоретически положительного эффекта, а воздействие в КС нестационарных помех затрудняет регулирование уровня порога в процессе приема.

Таким образом, методы повышения помехоустойчивости приема, основанные на использовании сигнала стирания, требуют решения задачи оптимизации уровня порога первой РС, т. е. минимизации вероятностей трансформации и ложного стирания при любых изменениях характеристик КС [4]. Математическая формулировка стоящей проблемы сводится к решению задачи автоматической классификации, состоящей в том, что необходимо отнести вектор $\mathbf{x}$ к одному из $\mathbf{s}$ классов множества $\{\mathbf{w_i}\}$, причем компоненты вектора $\mathbf{x}$ представляют собой "сведения" о наблюдаемом объекте. Простейшая нетривиальная математическая формулировка этой задачи возникает в случае, когда $\mathbf{s = 2}$ (ввиду дискретности сигнала, принимающего значения "0", либо "1") и линейная решающая функция

$$\mathbf{f(x) = w^T \cdot x + const}, \tag{1}$$

относит $\mathbf{x}$ к классу $\mathbf{w_1}$ при $\mathbf{f(x) > 0}$ и к классу $\mathbf{w_2}$ при $\mathbf{f(x) < 0}$ с допустимо малым числом неправильных классификаций. При этом процедура вычисления линейных решающих функций должна быть адаптивной (то есть реагировать на любые изменения помеховой обстановки в КС), быстро сходиться к локальному минимуму ошибки и выполняться без какого - либо априорного предположения относительно вида статистического распределения векторов $\mathbf{x}$ в каждом классе $\mathbf{w_i}$ [5]. Назовем ее "процедурой обучения с выровненными ошибками" или "ВО - процедурой". При этом слово "обучение" указывает на ее адаптивный характер.

Первые работы по непараметрическим обучаемым линейным классификаторам распадаются на две основные группы: процедуры коррекции ошибки [6] и процедуры минимизации квадратичной ошибки [7]. Однако, все эти процедуры оказались подвержены хотя бы одному из перечисленных недостатков: медленной сходимостью; сложностью вычислений и большим объемом требуемой для них памяти; плохой асимптотикой. Предложенная ВО - процедура лишена этих недостатков, поскольку [8]: она быстро сходится; требуется малый объем памяти (нет необходимости запоминать матрицы);значения асимптотики близки к минимуму вероятности ошибки.

Пусть $\mathbf{x}$ есть $\mathbf{d}$ - мерный вектор признаков $\mathbf{x = (x_1,...,x_d)}$, где $\mathbf{-\infty < x_i < \infty}$, и пусть $\mathbf{y}$ есть дополнительный вектор признаков $\mathbf{y = (x_0, x_1, ..., x_d)^T}$. Пусть $\mathbf{X(n)}$ и $\mathbf{Y(n)}$ есть векторные случайные переменные, являющиеся функциями номера шага $\mathbf{n}$ и порожденные обучающей последовательностью в соответствующих пространствах $\mathbf{x}$ и $\mathbf{y}$. В частности, $\mathbf{Y(n)}$ представляет собой дополнительный вектор признаков на шаге $\mathbf{n}$. Пусть $\mathbf{w(n)}$ есть класс, к которому принадлежит $\mathbf{Y(n)}$, и пусть $\{\mathbf{w_i}\}$ есть алфавит классов. Вследствие дискретности (двузначности) демодулируемого сигнала этот алфавит имеет длину два: $\{\mathbf{w}\}=\{\mathbf{w_1}, \mathbf{w_2}\}$. Пусть $\mathbf{w}$ есть весовой вектор

$(\mathbf{w}_1, \ldots, \mathbf{w}_d)^T$. Пусть $\mathbf{V(n)}$ есть векторная случайная переменная, порожденная процессом обучения в пространстве $\mathbf{v}$ и являющаяся также функцией номера шага.

Дополнительный вектор в процедуре обучения с пропорциональным приращением определяется рекурсивно следующим образом:

$$\mathbf{V(n+1) = V(n) + \rho_n \cdot Q(n)}, \tag{2}$$

где ρ_n – величина $\mathbf{n}$ - го шага и

$$\mathbf{Q(n)} = \begin{cases} \mathbf{Y(n)}, & \text{если } \mathbf{V(n)^T Y(n)} \quad \mathbf{w(n) = w_2}; \\ \mathbf{-Y(n)}, & \text{если } \mathbf{V(n)^T Y(n)} \quad \mathbf{w(n) = w_1}; \\ \mathbf{0}, & \text{в остальных случаях.} \end{cases} \tag{3}$$

В принципе, (2) есть разновидность метода градиентного спуска [7], при котором ищется минимум функции потерь $\mathbf{J(v)}$. Обычная процедура нахождения такого минимума сводится к определению рекурсивной функции $\mathbf{V(n)}$, сходящейся (стохастически) к нулю градиента функции потерь $\mathbf{J(v)}$. Такая рекурсивная функция может быть получена методами стохастической аппроксимации

$$\mathbf{J(v) = E(-v^T Q \mid V = v)}, \tag{4}$$

Получим для (4) другое выражение, приводящее непосредственно к ВО - процедуре обучения: пусть $\mathbf{f_i(x) = P(w_i)\ p(x|w_i)}$, где $\mathbf{P(w_i)}$ – априорная вероятность того, что $\mathbf{w = w_i}$ и $\mathbf{p(x|w_i)}$ – условная плотность распределения $\mathbf{X = x}$ при $\mathbf{w = w_i}$. Назовем $\mathbf{p(x|w_i)}$ плотностью, условной по классу, а $\mathbf{f_i(x)}$ – компонентной плотностью. Пусть ϑ_i есть область решения, соответствующая классу $\mathbf{w_i}$, тогда (4) можно записать в виде:

$$\mathbf{J(v) = |w| [M_1(v) + M_2(v)]}, \tag{5}$$

где $|\mathbf{w}|$ - длина вектора $\mathbf{w}$,

$$\mathbf{M_1(v) = P(w = w_1, y \in \vartheta_2) E\left(\frac{v^T y}{w} \middle| V = v, w = w_1, y \in \vartheta_2 \right)}, \tag{6}$$

$$\mathbf{M_2(v) = P(w = w_2, y \in \vartheta_1) E\left(\frac{-v^T y}{w} \middle| V = v, w = w_2, \lambda \in \vartheta_1 \right)}. \tag{7}$$

Легко показать, что $\dfrac{\mathbf{v^T y}}{|\mathbf{w}|}$ представляет собой расстояние между $\mathbf{x}$ и границей области ϑ_2, являющейся гиперплоскостью. Это расстояние положительно при $\mathbf{x} \in \vartheta_2$. Аналогично $\dfrac{\mathbf{-v^T y}}{|\mathbf{w}|}$ представляет собой расстояние между $\mathbf{x}$ и границей области ϑ_1, являющейся гиперплоскостью. Это расстояние положительно при $\mathbf{x} \in \vartheta_1$. Для случая двух классов обе границы совпадают. Таким образом, $\mathbf{M_i(v)}$ есть первый момент хвоста (ошибки) функции $\mathbf{f_i(x)}$.

Хотя эта процедура обучения асимптотически точна для линейно разделимых условных по классу плотностей $\{\mathbf{p(x|w_i)}\}$, её асимптотика может значительно отличаться от минимума вероятности ошибки в случае перекрывающихся плотностей, условных по классу. С другой стороны, минимум $\mathbf{M_1(v) + M_2(v)}$ имеет место при некотором значении $\mathbf{v}$, обозначаемом через $\mathbf{v_e}$, часто очень близком к обеспечивающему минимальную вероятность ошибки значению $\mathbf{v}$.

Обозначим через $\mathbf{v_p}$ значение $\mathbf{v}$, которое обеспечивает минимальную вероятность ошибки. Опыт [8] свидетельствует, что $\mathbf{v_e}$ и $\mathbf{v_p}$ часто очень близки друг к другу. Действительно, когда $\mathbf{f_1(x)}$ и $\mathbf{f_2(x)}$ симметричны относительно друг друга и, следовательно,

$$\mathbf{f_2(x) = f_1(b - x)}, \tag{8}$$

где $\mathbf{b}$ - центроид $\mathbf{f_1(x) + f_2(x)}$, тогда $\mathbf{v_e}$ и $\mathbf{v_p}$ совпадают.

Введение $|\mathbf{w}|$ в выражение для $\mathbf{J(v)}$ в качестве множителя (5) делает возможным существенное разделение точек минимумов функции $\mathbf{J(v)}$ и $\mathbf{v(p)}$, даже когда $\mathbf{v_e}$ и $\mathbf{v_p}$ совпадают. Использование (5) приводит также к смещению $\mathbf{W(n)}$ в сторону малых значений вектора $\mathbf{w}$ при $\mathbf{n\to\infty}$, поскольку $\mathbf{J(v) = 0}$ при $\mathbf{v = 0}$. В результате направление вектора $\mathbf{W(n)}$ при $\mathbf{n\to\infty}$ часто становится недостаточно определенным. Таким образом, асимптотическое поведение процедуры обучения с пропорциональным приращением может быть неудовлетворительным в случаях, когда плотности, условные по классу, перекрываются.

Для преодоления этого недостатка процедуры обучения с пропорциональным приращением используется функция потерь в виде

$$\mathbf{J(v) = M_1(v) + M_2(v)}, \tag{9}$$

что приводит к описываемой в статье ВО - процедуре обучения.

Получим ВО - процедуру обучения из функции потерь (9), используя методы градиентного спуска и стохастической аппроксимации. Предположим, что непрерывная дифференцируемая функция $\mathbf{J(v)}$ имеет единственный минимум при $\mathbf{v^*}$ и у нее отсутствуют локальные минимумы. Основная процедура градиентного спуска для такой функции $\mathbf{J(v)}$ представима в виде рекурсивного уравнения

$$\mathbf{v_{n+1} = v_n - \rho_n \nabla J(v_n)}, \tag{10}$$

где

$$\mathbf{\nabla J(v) = \text{градиент } J(v)}. \tag{11}$$

Тогда для любого достаточно малого $\mathbf{\rho_n}$ последовательность $\{\mathbf{J(v^*)}\}$ будет монотонно убывающей и сходящейся при $\mathbf{n\to\infty}$ к $\mathbf{J(v^*)}$. Метод градиентного спуска был расширен [8] на случай «зашумленных» функций, т. е. функций $\mathbf{J(v)}$, зависящих от одной или нескольких случайных переменных. Такое расширение приводит к стохастической аппроксимации. Стохастическая аппроксимация связана с нахождением экстремумов и нулей регрессивной функции, т. е. среднего значения случайной переменной как функции, зависящей от параметра. Предположим, что $\mathbf{J(v)}$ является регрессивной функцией и что, кроме того, можно найти такую случайную переменную $\mathbf{Z}$, зависящую от $\mathbf{v}$, что

$$\mathbf{E(Z|\nu) = -\nabla J(\nu_n)}\,. \tag{12}$$

Также предположим, что $\mathbf{J(v)}$ имеет единственный минимум в точке $\mathbf{v}^*$. Если бы $\mathbf{E(Z|v_n)}$ было известно, то $\mathbf{v}^*$ можно было бы найти при помощи градиентного спуска (10):

$$\mathbf{v_{n+1} = v_n + \rho_n E(Z|v_n)}\,. \tag{13}$$

Обычно при малых $\mathbf{n}$ $\mathbf{E(Z|v_n)}$ неизвестно. В этом случае при методе стохастической аппроксимации в выражении (13) используется замена $\mathbf{E(Z|v_n)}$ на полученное, на шаге $\mathbf{n}$ значение $\mathbf{Z}$, то есть $\mathbf{Z(n)}$. При этом условии $\mathbf{v_n}$ становится случайной переменной, которую обозначим через $\mathbf{V(n)}$. Таким образом, (13) преобразуется к виду

$$\mathbf{V(n+1) = V(n) + \rho_n Z(n)}. \tag{14}$$

Из теории стохастической аппроксимации следует, что стохастическая сходимость $\{\mathbf{V(n)}\}$ к $\mathbf{v}^*$ зависит от выбора $\boldsymbol{\rho_n}$, $\mathbf{Z(n)}$ и регистрационных функций. Например, если $\boldsymbol{\rho_n}$ уменьшается слишком быстро при увеличении $\mathbf{n}$, то $\mathbf{V(n)}$ не сходится к $\mathbf{v}^*$.

Имея в виду, что процедура обучения работает в дополнительном весовом пространстве, будем считать $\mathbf{V(n)}$ в (14) дополнительным весовым вектором на шаге $\mathbf{n}$ и $\mathbf{Z(n)}$ соответствующим образом выбранной функцией, зависящей от $\mathbf{Y(n)}$, $\mathbf{V(n)}$ и $\mathbf{w(n)}$. Для соответствующего выбора $\mathbf{Z(n)}$ в случае обычной стохастической аппроксимации пришлось бы искать случайную переменную $\mathbf{Z}$, удовлетворяющую (12). Однако, выбор $\mathbf{Z}$ требует особой осторожности в случае, когда $\mathbf{J(v)}$ является функцией потерь для ВО - процедуры обучения (9), потому что в этом случае $\mathbf{J(v)}$ не зависит от $|\mathbf{v}|$. Эта независимость приводит к неоднозначности положения минимума функции $\mathbf{J(v)}$. В самом деле, если $\mathbf{J(v)}$ является минимумом в точке $\mathbf{v}$, то это также минимум и при $\mathbf{v}^*$ для всех действительных значений $\mathbf{a}$. При этом $\mathbf{Z}$ должна быть выбрана такой, чтобы свойства сходимости $\mathbf{V(n)}$ не зависели от $|\mathbf{V(n)}|$, а зависели только от $\mathbf{V(n)/|V(n)|}$.

Описанная выше ВО - процедура была применена для систем связи с целью оптимизации уровня порога первых решающих схем (РС) двоичных стирающих каналов (ДСК) путем выработки на основе ВО - процедуры управляющего воздействия, пропорционального изменению входных параметров демодулируемого сигнала [8]. При этом минимизируются вероятности трансформации и ложного стирания символа при изменении помеховой обстановки в КС, чем повышается помехоустойчивость приема в целом.

Примером технической реализации может служить устройство для адаптивного мажоритарного декодирования телемеханических дублированных сигналов. Структурная схема устройства, реализующего ВО - процедуру обучения, задаваемую уравнениями (2) и (3), приведена на рис.1.

Здесь математическое выражение (3) выглядят следующим образом:

$$\mathbf{Q(n)} = \begin{cases} \mathbf{Y(n), V(n)^T Y_n < \Pi} \\ \mathbf{-Y(n), V(n)^T Y_n > \Pi} \end{cases}, \tag{15}$$

где **V(n)** - коэффициент усиления усилителя 1 с регулируемым коэффициентом усиления (УРКУ); ρ_n - шаг, с которым данная процедура сходится к минимуму вероятности ошибки и реализуется коэффициентом усиления усилителя 4 с постоянным коэффициентом усиления (УПКУ), выставляемым вручную или автоматически в зависимости от характера нестационарной помехи в канале связи; **Y(n)** - образцы канальных сигналов искаженных помехой и поступающие на вход данного устройства (анализатора помех); **П** - величина порога блока 3 и сходимость которой к минимуму вероятности ошибки доказана.

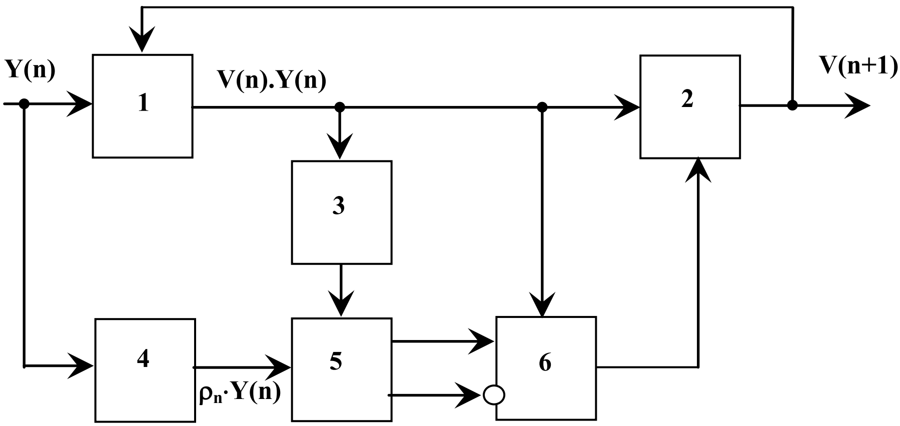

Рис. 1. Структурная схема устройства, реализующего ВО - процедуру обучения

Устройство работает следующим образом. Входной сигнал **Y(n)** поступает на входы УРКУ (блок 1) и УПКУ (блок 4). Коэффициент усиления УРКУ имеет в различные моменты времени значения **V(n)**, **V(n+1)**, при этом УРКУ реализует действие **V(n)·Y(n)**. УПКУ имеет коэффициент усиления ρ_n , определяющий шаг схождения процедуры, и реализует действие $\rho_n \cdot$**Y(n)**. Сигнал с выхода УРКУ поступает на вход компаратора (блок 3), который производит анализ выполнения условий **V(n)Y(n) < 0**, либо **V(n)Y(n) > 0**. При выполнении первого неравенства переключатель (блок 5), управляемый сигналом с выхода компаратора, пропускает усиленный сигнал с выхода УПКУ реально, либо инверсно при выполнении второго неравенства. Таким образом, реализуется условие $\pm\rho_n\mathbf{Y(n)}$. Для общего случая компаратор проводит сравнение не с уровнем логического нуля, а с опорным напряжением, первоначально устанавливаемым при подаче на вход устройства сигнала при отсутствии помех в КС. Сумматор (блок 6) выполняет действие $\mathbf{V(n)} \pm \rho_n\mathbf{Y(n)}$. После этого полученное значение $\mathbf{V(n+1)}$ сравнивается в блоке сравнений (вычитатель) (блок 2) со значением $\mathbf{V(n)}$. Сигнал несоответствия, пропорциональный разности сравниваемых величин, поступает на вход УРКУ по обратной связи ОС и регулирует коэффициент его усиления, чем реализуется зависимость $\mathbf{V(n+1) = f[V(n)]}$. Кроме того, полученное управляющее воздействие есть искомая процедура ВО - обучения, которой присущи все описанные ранее достоинства.

По указанному алгоритму в зависимости от уровня помех в канале связи на выходе устройства вырабатывается сигнал, уровень которого пропорционален изменению помехи, действующей в данный момент времени в канале связи. Этот сигнал регулирует ширину полосы "стирания" детектора качества.

Для оптимизации уровня порога первых решающих схем следует применять математический метод градиентного спуска, который приводит к процедуре обучения с выравненными ошибками. Простота технической реализации данной процедуры позволяет применять её в первых решающих схемах СПДИ без их значительного усложнения и увеличения времени принятия решения.

ЛИТЕРАТУРА

1. Берлекэмп Э.Р. Техника кодирования с исправлением ошибок. ТИИЭР. – 1980. - Т.68, № 5. - С. 24 - 58.

2. Обнаружение и исправление ошибок в дискретных устройствах./ Под. ред. В.С. Толстякова. – М.: Сов. радио, 1972. - 297 с.

3. Шувалов В.П. Косвенные методы обнаружения ошибок. – М.: Связь, 1972. - 362 с.

4. Камыш А.В. Способ дельта-модуляции / Сб. тезисов IX НТК СВВИУС 1995 г. – Ставрополь: СВВИУС, 1995. - С. 73 - 74.

6. Rosenblatt F. The perceptron: a probabilistic model for information storage and organisation in the brain. - 1958. - Vol. 65, № 6. P. 386 - 408.

7. Ho Y.C., Kashyap R.L. An algorithm for linear inequalities and its applications // IEEE Trans. El. Comp. – 1965. - Vol. 14, № 10. – P. 683 - 688.

8. Вассель Г.Н., Склански Дж. Адаптивный непараметрический классификатор // ТИИЭР. – 1976. - № 8. - С. 52 - 62.