УДК 004.931 (045)

Яценко Н.Н.

РАСПОЗНАВАНИЕ ТРАФИКА БЕСПРОВОДНОЙ КОМПЬЮТЕРНОЙ СЕТИ НА ФИЗИЧЕСКОМ УРОВНЕ

Национальный авиационный университет

Предложен метод анализа трафика беспроводной компьютерной сети на физическом уровне. Данные передаваемые рассматриваются в виде синусоиды для последующего разложения в ряд Фурье и анализа полученных данных.

Актуальность темы

Постоянно возрастающее количество передаваемых данных через общедоступные сети вызывает необходимость проведения анализа трафика для разнообразных целей:

- обнаружение несанкционированных действий в сети;
- исследование работы сети для последующего усовершенствования;
- обнаружение неисправных участков сети для устранения неисправностей и т.д.

Таким образом, возникает необходимость в построении специальных систем распознавания. Необходимость системного подхода к проблеме распознавания обуславливается рядом причин. Первая состоит в том, что распознавание не является самоцелью. Она является всего лишь средством получения информации, необходимой системе управления для выработки определенного решения, стратегии поведения или стратегии управления. Следовательно, система распознавания должна строится таким образом, чтобы обеспечивалась наибольшая эффективность системы управления, стоящей над системой распознавания. Это означает подчиненность цели системы распознавания целям системы управления.

Вторая причина состоит в том, что эффективность системы распознавания в целом непосредственно зависит от эффективности технических средств системы распознавания (измерительных и вычислительных) и ее математического обеспечения – программно реализованных алгоритмов построения описания классов объектов на языке признаков, обработки измерительной информации в целях определения признаков, собственно распознавания, корректировки априорных описаний и т.д. Это в свою очередь означает подчиненность целей средств системы распознавания цели системы в целом [1].

Постановка задачи

Наиболее распространенной абстрактной моделью распознавания образов является модель классификации. Модель эта состоит из трех частей — датчика, выделителя признаков и классификатора. Датчик воспринимает воздействие и преобразует его к виду, удобному для машинной обработки. Выделитель признаков (называемый также рецептором, фильтром свойств, детектором признаков или препроцессором) выделяет из входных данных предположительно относящуюся к делу информацию. Классификатор на основе этой информации относит эти данные к одной из нескольких категорий. Обсуждение принципа действия или конструкции датчика не входит в задачи данной статьи. Выделение признаков и классификация, напротив, представляет для нас интерес. С точки зрения теории провести черту между этими двумя действиями можно весьма условно. Идеальный выделитель признаков предельно упрощает работу классификатора, тогда как при наличии всемогущего классификатора не требуется помощь выделителя признаков. Различие это хотя и весьма важно, но делается оно, исходя скорее из практических, нежели теоретических соображений.

Задача выделения признаков более специализирована по сравнению с задачей классификации. Хороший выделитель признаков, предназначенный для одной цели, принесет, по-видимому, малую пользу при классификации других объектов. Однако большое число технических приемов было развито в связи с задачей извлечения полезной информации из изображения. [2]

Задача классификации по существу представляет собой задачу разбиения пространства признаков на области, по одной для каждого класса. Разбиение это в принципе надо производить так, чтобы не было ошибочных

решений. Если этого сделать нельзя, то желательно уменьшить вероятность ошибки, или если ошибки имеют различную цену, то сделать минимальной среднюю цену ошибки. При этом задача классификации превращается в задачу статистической теории принятия решений, широко применяемую в различных областях теории распознавания образов.

Таким образом первой задачей при разработке системы распознавания является конструирование выделителя признаков.

Основная часть

Реальные объекты обладают бесконечным числом признаков. С другой стороны, классификация, производимая человеком, обычно основывается на небольшом числе признаков, как например, максимальная величина, основная частота и т. д. Каждое из этих измерений несет значительную информацию для целей классификации и выбирается в соответствии с физическим смыслом задачи.

Очевидно, что с уменьшением числа входных величин классификатора его проектирование упрощается. Для того чтобы добиться этого, следует наметить некоторые пути для выбора или извлечения существенных информативных признаков из всей совокупности наблюдаемых. Эту задачу называют задачей выбора информативных признаков, и она составляет другой важный раздел теории распознавания образов. [3] Выбор признаков можно рассматривать как отображение исходного n-мерного пространства в пространство меньшей размерности. При этом необходимо сохранить свойство разделимости распределений, соответствующих разным классам. Следовательно, отображение должно быть выполнено без существенной потери этого свойства.

Таким образом, как показано на рис. 1, задача распознавания образов состоит из двух частей: выбор информативных признаков и проектирование классификатора. На практике между этими частями нет четкой границы. Действительно, классификатор можно представить как устройство для выбора признаков, которое отображает m признаков в один (дискриминантная функция). Однако при проектировании системы распознавания удобнее разделить задачу распознавания на две части и изучать их независимо друг от друга.

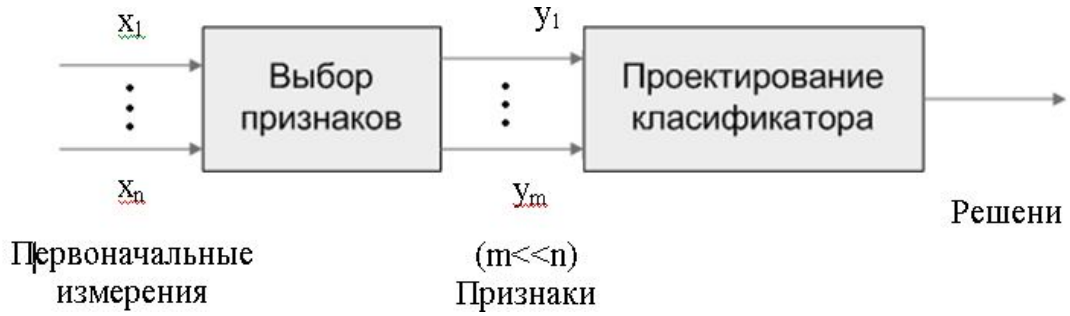

Рис.1 Блок-схема системы распознавания образом

При распознавании реальных объектов обычно возникает задача представления модели данного объекта. Трафик на физическом уровне в беспроводной компьютерной сети имеет вид синусоиды. Поэтому наиболее логичным является представление его модели в виде функции описывающей график трафика. Для имитации данного подхода было разработано приложение позволяющее моделировать трафик передаваемый в беспроводной компьютерной сети. Графический интерфейс данного приложения показан на рис. 2.

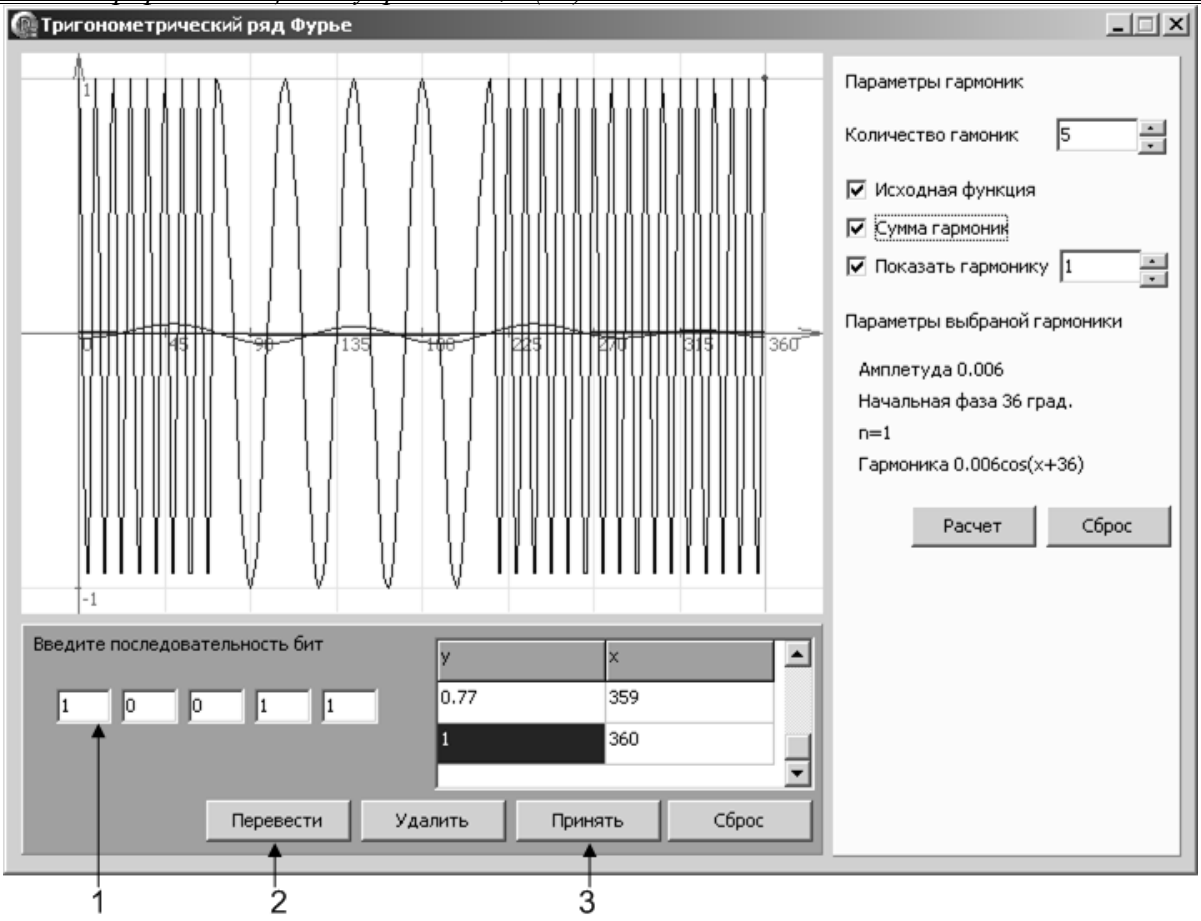

Рис. 2 Графический интерфейс приложения моделирующего трафик

Принцип работы разработанного приложения

На рис. 2 цифрой 1 обозначен элемент программы, используемый для ввода тестовой последовательности бит. Цифрой 2 обозначена кнопка «Принять», после нажатия которой вычисляются координаты последовательности точек графика моделирующего последовательность бит в беспроводной компьютерной сети. Для кодирования бит используется частотная модуляция.

Цифрой 3 обозначена кнопка «Прощет», которая инициирует разложение в ряд Фурье, полученной ранее функции. Поскольку задан интервал разложения $-\pi < t < \pi$, то ряд Фурье, порожденный действительной функцией $f(t)$, для которой существует интеграл $\int_{-\pi}^{\pi} |f(t)|dt$, есть бесконечный тригонометрический ряд

$$\frac{1}{2}a_0 + \sum_{k=1}^{\infty}\left(a_k \cos kt + b_k \sin kt\right) \equiv \sum_{k=-\infty}^{\infty} c_k e^{ikt}$$

Коэффициенты которого определяются по формулам Эйлера-Фурье

$$\left.\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos kt\ dt,\ \ b_k = \frac{1}{\pi}\int_{-\lambda}^{\pi} f(t)\sin kt\ dt, \\ c_k &= \bar{c}_{-k} = \frac{1}{2}\left(a_k - ib_k\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}dt \end{aligned}\right\} (k = 0,1,2...).$$

В разделе «Параметры гармоник» задается количество гармоник и управление отображаемыми данными на графике для уменьшения загруженности. В разделе «Параметры выбранной гармоники» отображаются данные для последующего анализа трафика. Это такие параметры функции как амплитуда, частота и фаза.

Выводы

Данный метод моделирования передаваемого трафика в беспроводной сети дает воз-

можность проводить анализ не распаковывая фрейм на канальном уровне, пакет на сетевом или сегмент на канальном. Это повышает скорость обработки данных, т.к. общеизвестным фактом является то, что чем ниже уровень модели OSI на котором работает сетевое устройство тем меньшую задержку при передаче данных оно создает. Следующим этапом проектирования системы распознавания трафика в беспроводной компьютерной сети является анализ наиболее информативных признаков с целью последующей классификации распознаваемого объекта.

В дальнейшем необходимо будет определить критерий качества идентификации. В подавляющем большинстве работ он выбирался квадратичным в виде среднего значения квадрата невязки [4]. Минимизация такого квадратичного критерия во многих случаях сводится к решению системы алгебраических уравнения. Возможность получения теоретически точного результата на основе различных вариантов метода наименьших квадратов обеспечила господствующее положение квадратичному критерию идентификации. Значительно реже используются критерии качества идентификации, отличные от квадратичных, например, модульный критерий типа среднего значения абсолютной величины невязки.

Литература

1. А.Л. Горлкин, И.Б. Гуревич, В.А. Скрипник, Современное состояние проблемы распознавания: некоторые аспекты. – М.: Радио и связь, 1985. – 160 с.

2. Р. Дуда, П. Харт, Распознавание образов и анализ сцен. – М.: Мир, 1976. – 507 с.

3. К. Фукунага, Введение в статистическую теорию распознавания образов. – М.: Наука. Главная редакция физико-математической литературы, 1979. – 368 с.

4. Я.З. Цыпкин, Информационная теория идентификации. – М.: Наука. Физматлит, 1995. – 336 с.