

Оценка информационной эффективности устройств цифровой обработки информации

А. Н. Губин

Санкт-Петербургский государственный
университет телекоммуникаций
им. проф. М. А. Бонч-Бруевича
Санкт-Петербург, Россия
gan50_60@mail.ru

В. Л. Литвинов

Санкт-Петербургский государственный
электротехнический университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
Санкт-Петербург, Россия
vlad.litvinov61@gmail.com

Ф. В. Филиппов

Санкт-Петербургский государственный
университет телекоммуникаций им. проф. М. А. Бонч-Бруевича
Санкт-Петербург, Россия
9000096@mail.ru

Аннотация. Рассматривается широкий класс устройств цифровой обработки информации – от классических фильтров до нейронных сетей. Основное назначение рассматриваемого класса объектов – извлечение нужной информации из входных сигналов. В докладе предлагается подход к количественной оценке информационной эффективности преобразователей информации, основанный на анализе изменений энтропии, понятии Колмогоровской сложности и расстояния Кульбака-Лейблера.

Ключевые слова: информационная эффективность; энтропия; Колмогоровская сложность; расстояние Кульбака-Лейблера

I. ВВЕДЕНИЕ

Для извлечения информации из различных данных используется большое многообразие устройств цифровой обработки. При этом решается широкий класс задач от классической фильтрации сигналов до распознавания образов и генерации цифровых объектов с заданными свойствами. В любом случае результат извлечения нужной информации можно характеризовать достигнутым качеством решения поставленной задачи и привлеченными затратами. Именно эти две характеристики являются ключевыми для оценки информационной эффективности того или иного использованного устройства. Попробуем отыскать подход к количественной оценке информационной эффективности универсальных устройств цифровой обработки информации.

II. ПОСТАНОВКА ЗАДАЧИ

Процесс извлечения полезной информации из входных данных базируется на использовании априорной информации как о статистических характеристиках входного множества, так и характеристиках подлежащих воспроизведению выходных данных. Причем, очевидно, что чем больше такой информации используется при

реализации цифровых устройств, тем эффективней решается задача извлечения полезных данных. В общем случае, соотношения между взаимной информацией и условными Колмогоровскими сложностями (КС) [1] при взаимодействии объектов (входной сигнал – цифровое устройство – выходной сигнал) можно наглядно представить в виде схемы (рис. 1).

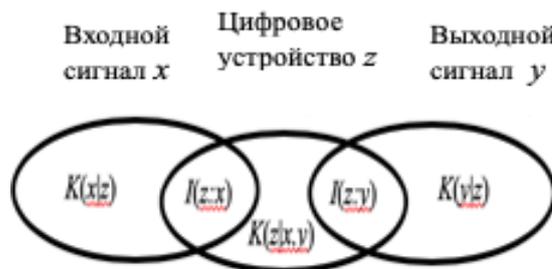


Рис. 1. Общая схема информационного взаимодействия

На рис. 1 область $I(z|x)$ представляет собой количество информации в устройстве z о входном сигнале x , которое использовано при реализации и настройке устройства для обработки этого сигнала. Аналогично $I(z|y)$ обозначает количество информации в устройстве z о выходном сигнале y . Под $K(x|z)$ понимается условная Колмогоровская сложность входного сигнала x при его взаимодействии с цифровым устройством z , $K(y|z)$ означает условную КС выходного сигнала y , а $K(z|x,y)$ – условная КС цифрового устройства, формирующего сигнал y при обработке входного сигнала x .

Очевидно, что чем больше априорной информации о параметрах входного сигнала x и подлежащего выделению из помех сигнала y используется при определении параметров цифрового устройства z , тем эффективнее будет решаться задача выделения выходного сигнала y ,

следовательно, тем большие значения $I(z: x)$ и $I(z: y)$ будут выделены из величины $K(z)$.

Таким образом, для оценки информационной эффективности цифровых устройств можно использовать коэффициент информационной эффективности $K_{из}$:

$$K_{из} = \frac{I(z: x) + I(z: y)}{K(z)}$$

где $K(z)$ – простая Колмогоровская сложность цифрового устройства.

III. ИНФОРМАЦИОННАЯ ЭФФЕКТИВНОСТЬ ЦИФРОВЫХ ФИЛЬТРОВ

Классическим представителем устройств цифровой обработки информации является сглаживающий фильтр. Рассмотрим процесс сглаживания случайных сигналов рекурсивными цифровыми фильтрами, которые обеспечивают несмещенное воспроизведение регулярной составляющей входного сигнала и подавление аддитивной стационарной помехи с заданной корреляционной функцией. Если входной сигнал сглаживающего устройства может принимать n дискретных значений x_i ($i = 1, 2, \dots, n$) с некоторой вероятностью $p(x)$, то неопределенность состояний x_i выражается энтропией:

$$H_x = - \sum_{k=1}^n p(x_k) \log p(x_k).$$

Аналогично можно записать выражение для энтропии выходного сигнала y . Процесс сглаживания помех и несмещенное воспроизведение полезного сигнала базируется на том, что при определении параметров передаточной функции цифрового фильтра используется априорная информация о характеристиках помех и регулярной составляющей входного и выходного сигналов.

Сумма $I(z: x) + I(z: y)$ представляет собой общее количество априорной информации, используемое при определении параметров сглаживающего фильтра. Состояние неопределенности для процесса сглаживания случайных сигналов можно охарактеризовать значениями энтропийных потенциалов входного Δ_{ex} и выходного Δ_{ey} сигналов, а динамику изменения состояния неопределенности – отношением энтропийных потенциалов:

$$\frac{\Delta_{ex}}{\Delta_{ey}} = \frac{\frac{1}{2} e^{H(x)}}{\frac{1}{2} e^{H(y)}} = e^{H(x)-H(y)} = e^{I(x,y)},$$

где $I(x, y)$ – количество информации, порожденное процессом сглаживания случайного сигнала x .

Отсюда получаем:

$$I(x, y) = \ln \frac{\Delta_{ex}}{\Delta_{ey}} = \ln \frac{K_{ex} \sigma_x}{K_{ey} \sigma_y} = \ln \frac{K_{ex}}{K_{ey}} + \ln \frac{\sigma_x}{\sigma_y} = I_i + I_p,$$

где K_e – значение энтропийного коэффициента, а σ – среднеквадратическое отклонение соответствующего сигнала.

Здесь I_i – составляющая, величина которой определяется законом распределения вероятностей значений входного и выходного сигналов, а I_p – энергетическая составляющая количества информации. Очевидно, I_i при вычислении $K_{из}$ можно исключить, так как эффективность фильтрации оценивается для различных фильтров на примере работы с одинаковыми сигналами. Таким образом, окончательное выражение для коэффициента информационной эффективности фильтра определится как:

$$K_{из} = \frac{\ln \frac{\sigma_x}{\sigma_y}}{K(z)},$$

где Колмогоровская сложность фильтра может рассматриваться как минимальная длина программы, реализующая фильтр z при наличии информации о x и y [2].

IV. ИНФОРМАЦИОННАЯ ЭФФЕКТИВНОСТЬ АВТОЭНКODЕРОВ

В качестве классической модели цифровых преобразователей класса нейронных сетей рассмотрим автокодировщики (автоэнкодеры). Основной принцип работы и обучения автокодировщика – получить на выходном слое отклик, наиболее близкий к входному. С их помощью решается широкий спектр задач, и они представляют, пожалуй, лучший и наиболее адекватный образец для оценки информационной эффективности на базе Колмогоровской сложности.

Действительно, в [3] отмечалось, что на базе автокодировщиков можно получать достаточно точные оценки Колмогоровской сложности, которые можно использовать для оценки информационной эффективности.

Для нашей цели используем общую модель автокодировщика, который реализует последовательное применение функции кодера $h = g(x)$ и декодера $y = f(h)$, где x – входной объект, h – его латентное представление или сжатый образ, а y – восстановленный входной объект. Общая структура автокодировщика показана на рис. 2.

Практические примеры использования подобной модели позволяют объединить вероятностные определения энтропии, использованные выше, с алгебраической концепцией, введенной Колмогоровым [4]. Например, при решении задачи понижения размерности, сжатый образ h входного объекта x можно использовать для оценки энтропии:

$$H(y|h) = \min_{P(h)=y} l(P),$$

как «программы» P минимальной длины, которая позволяет получить объект y из объекта h . Очевидно, в

этом случае в качестве КС может выступать значение $l(h)$, которое может служить оценкой оптимальности латентного представления входного объекта.

При этом, качество реконструкции объекта y из h определяется значением достигнутой среднеквадратичной ошибки $L(x, f(g(x)))$, штрафующей выходной вектор y за отличие от x .

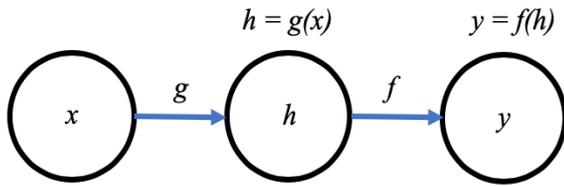


Рис. 2. Общая модель автокодировщика

Если перейти к вероятностному подходу, когда качество работы преобразователя оценивается на основе анализа входного P и выходного Q вероятностного распределения, удобно использовать расстояние (дивергенцию) Кульбака-Лейблера:

$$KL(P||Q) = \sum_{i=1}^n p(x_i) \log \frac{p(x_i)}{q(x_i)}.$$

В теории информации оно является количеством информации, которая теряется при приближении распределения P с помощью распределения Q . Расстояние Кульбака-Лейблера не является метрикой в пространстве вероятностных распределений, так как оно не симметрично и не удовлетворяет неравенству треугольника. При этом расстояние Кульбака-Лейблера всегда неотрицательно и равно нулю только тогда, когда распределения P и Q совпадают почти всюду. Таким образом, будучи предварительной метрикой, оно порождает топологию в пространстве вероятностных распределений.

Доказано [5], что расстояние Кульбака-Лейблера – это единственная мера разницы между вероятностными распределениями, которая удовлетворяет некоторым желательным свойствам, являющимся каноническими расширениями в часто используемых характеристиках энтропии.

Один из способов интерпретации оценки качества реконструкции состоит в том, чтобы рассматривать ее как задачу минимизации расстояния Кульбака-Лейблера. Поскольку $p(x_i)$ зависит только от обучающего набора данных, то при обучении модели мы изменяем только величину $q(x_i)$. По сути, минимизация расстояния Кульбака-Лейблера в точности соответствует минимизации перекрестной энтропии между распределениями [6].

Таким образом, расстояние Кульбака-Лейблера позволяет оценить общее количество информации,

задействованное в процессе обработки входных векторов из входного вероятностного распределения P и перекрестную энтропию $H(p, q)$. Для этого предыдущее выражение можно преобразовать к следующему виду:

$$KL(P||Q) = \sum_{i=1}^n p(x_i) \log p(x_i) - \sum_{i=1}^n p(x_i) \log q(x_i) = H(p) + H(p, q).$$

Поскольку анализ информационной эффективности предполагает оценивать «прирост» количества информации, то коэффициент для оценки информационной эффективности можно предложить в виде:

$$K_{из} = \frac{KL(P||Q) - H(P)}{K(h)},$$

где числитель оценивает перекрестную энтропию, которая достигнута в результате обучения автокодировщика, а знаменатель характеризует минимальные затраты на реализацию модели. Как подчеркивалось в работе авторов [3], для оценки Колмогоровской сложности автокодировщика следует брать не программу моделирования автокодировщика, а результирующую нейронную сеть, полученную после обучения.

V. ЗАКЛЮЧЕНИЕ

В заключение отметим, что предложенные подходы к оценке информационной эффективности могут быть полезны при построении цифровых устройств обработки информации и выбора наиболее эффективного варианта практической реализации.

СПИСОК ЛИТЕРАТУРЫ

- [1] Колмогоров А.Н. Три подхода к определению понятия "количество информации" // Проблемы передачи информации. 1965. Том 1. Выпуск № 1. С. 3-11.
- [2] Верещагин Н.К. Колмогоровская сложность и алгоритмическая случайность / Н.К. Верещагин, В.А.Успенский, А. Шень. М.: МЦНМО, 2013. 576 С.
- [3] Губин А.Н., Литвинов В.Л., Литвинов Д.В., Филиппов Ф.В. Измерение колмогоровской сложности двоичных строк на базе автоэнкодеров. // Актуальные проблемы инфотелекоммуникаций в науке и образовании. VIII Международная научно-техническая и научно-методическая конференция; сб. науч. ст. в 4 т. / Под. ред. С.В. Бачевского; сост. А.Г. Владыко, Е.А. Аникевич. СПб. : СПбГУТ, 2019. Т. 2. С. 221-224.
- [4] Li M., Vitanyi P. An introduction to Kolmogorov complexity and its applications. Springer Verlag. Third Edition, 2008. ISBN 987-0-387-49820-1.
- [5] Hobson Arthur. Concepts in statistical mechanics. Gordon and Breach. New York, 1971. — ISBN 0677032404.
- [6] Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А.А. Слинкина. 2-е изд., испр. М.: ДМК Пресс, 2018. 652 с.