

Использование нейросетей для прогнозирования превышения заданного порога длинной очереди сообщений, отправляемых по скоростной магистрали

О.Д. Бербасов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
olegoleg.ru@mail.ru

А.Ю. Привалов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
privalov1967@gmail.com

Аннотация—В работе предложен подход к предсказанию превышения заданного порога значением длины очереди пакетов трафика, отправляемых по высокоскоростной магистрали.

Ключевые слова— прогнозирование, очередь, трафик, нейросети.

1. ВВЕДЕНИЕ

Производительность компьютерной сети является наиболее значимым параметром как для поставщика услуг связи, так и для потребителей их услуг. При проектировании таких сетей важно учитывать множество параметров, одним из которых является нагрузка. Нагрузка это один из параметров, которые влияют на отказоустойчивость и стабильность работы сети. Использование моделирования на этапе проектирования позволяет оценить оптимальную конфигурацию сети и нагрузку на неё в процессе работы, благодаря чему становится проще прогнозирование отказов и оценка её дальнейшей модернизации.

Для создания адекватных моделей необходимо анализировать реальные записи трафика, собранные в различных условиях. При этом, обнаруживается, что реальный трафик обладает особенностями, которые не встречаются у классических моделей теории телетрафика [1,2].

Подобными исследованиями и сбором данных занимаются различные организации по всему миру [2-4], одним из которых является CAIDA [5]. Центр прикладного анализа интернет-данных (CAIDA) проводит сетевое исследование и создаёт исследовательскую инфраструктуру для поддержки крупномасштабного сбора, обработки и распространения данных в научно-исследовательском сообществе. Данные, полученные с этого ресурса, используются для анализа характеристик реальных трасс и для обучения нейронных сетей.

2. ОСНОВНАЯ ЧАСТЬ

В работе используются записи трафика из базы CAID, которые записывались в 2016 году в Чикагском дата-центре Equinix. В процессе обработки сырых данных, отсеиваются все данные, кроме временных меток и значений полезной нагрузки сообщений на уровне IP-пакета, а также время, в течении которого группируются сообщения.

По этим данным были сформированы выборки, которые состоят из временных отметок и суммы объемов полезной нагрузки, переданных за это время сообщений.

Для каждой выборки были рассчитаны выборочное среднее, выборочная средняя дисперсия, корреляционная и автокорреляционные функции. Эти данные в дальнейшем будут использоваться в процессе подготовки обучающих наборов для нейронной сети.

Перед созданием обучающих наборов, требуется преобразовывать записи трафика в запись результатов моделирования ухода сообщений из дискретной очереди. Значимой величиной здесь будет являться объём данных в очереди в каждый момент времени.

Основным параметром для построения модели дискретной очереди будут являться средний уход из очереди и значения входного потока в каждый момент времени. При построении модели дискретной очереди средний уход C будем определять следующим образом:

$$C = MX / k, \quad (1)$$

где C – средний уход, k – коэффициент загрузки, а MX – выборочное среднее входного потока.

Моделируя процесс ухода сообщений из очереди получим значение длины очереди в каждый момент времени. Полученная выборка значений длин очереди будет использоваться для формирования обучающих примеров. В данной работе, значение порога, превышение которого будет предсказывать нейронная сеть, берется равным выборочному среднему этой выборки.

3. ИСПОЛЬЗОВАНИЕ НЕЙРОННОЙ СЕТИ

Для решения задачи, описанной выше, были рассмотрены различные виды нейронных сетей, которые по существующему опыту применения, чаще всего рекомендуются для решения задач прогнозирования.

В работе используется нейронная сеть, состоящая из четырех слоёв:

- первый слой (входной) который, содержит столько нейронов, сколько отсчетов входит в окно прогнозирования;
- второй слой (скрытый) – это LSTM [6] слой, содержащий 64 нейрона, использующих в качестве функции активации гиперболический тангенс;

- третий слой (скрытый) – слой пакетной нормализации [7];
- четвёртый слой (выходной) – слой с единственным нейроном, с линейной активацией.

В качестве метода оптимизации применяется метод стохастического градиентного спуска Adam.

4. ФОРМИРОВАНИЕ ОБУЧАЮЩИХ НАБОРОВ

Для обучения нейронной сети, её необходимо обучить. В данной работе используются методы обучения нейронных сетей с учителем. Для этого используются обучающие примеры, каждый из которых содержит некоторый набор данных.

Содержимым обучающих примеров является набор из входных и выходных данных нейронной сети. Входные данные являются одномерным вектором, который содержит значения длины очереди во временном окне заданной ширины. Выходные данные содержат значения длины очереди, в момент времени, следующий за соответствующим временным окном. Значение ширины окна, промежуток времени до заданного значения, а также параметры нейронной сети – это значения, которые подбираются в процессе анализа прогнозов, полученных от обученной нейронной сети и её переобучении.

5. ОБРАБОТКА РЕЗУЛЬТАТОВ РАБОТЫ НЕЙРОННОЙ СЕТИ

В процессе прогнозирования, обученная нейронная сеть, выдает значения длины очереди в момент времени, следующий за поданным на её вход окном прогнозирования. Таким образом, следующим шагом, является сравнение, полученного от нейронной сети прогноза с некоторым заданным пороговым значением. Результат этого сравнения будет являться результатом прогноза разработанной модели.

6. РАЗРАБОТКА СИСТЕМЫ

На данном этапе, разрабатываемая система прогнозирования превышения длиной очереди сообщений этого порога через заданный промежуток времени, состоит из двух подсистем: подсистема обработки данных и подсистема построения нейронных сетей и обработки результатов их вычислений.

Подсистема обработки данных разрабатывается на языке программирования C#. Она обеспечивает предобработку «сырых» данных файлов *.pcap [6] и дальнейшую работу с этими данными.

В процессе работы этой подсистемы, по исходным данным формируются выборки, которые анализируются и с учетом указанных параметров, преобразовываются в данные для построения модели дискретной очереди.

Данная подсистема так же обеспечивает моделирование процесса ухода сообщений из дискретной очереди по заданным параметрам. Таким образом строится файл данных, содержащий информацию о состоянии очереди и загрузке канала в каждый момент времени.

После формирования файла с данными очередей, подсистема строит обучающие наборы, состоящие из обучающих примеров. Процесс построения обучающих примеров на данном этапе работы часто меняется. Здесь,

по заданным параметрам, формируются обучающие примеры, структуру которых можно задавать через параметры системы.

Подсистема работы с нейронными сетями разрабатывается на языке программирования Python, с использованием технологий tensorflow и keras. Так как структуру обучающих наборов приходится часто менять, был написан специальный адаптер, который строит нейронную сеть используя указанные в подсистеме параметры и данные о размерностях, которые можно извлечь из обучающих наборов.

Результатом работы подсистемы работы с нейронными сетями является обученная нейронная сеть, с помощью которой можно получать прогноз по некоторому окну данных и порогу. Остальные параметры, такие как время до прогнозируемого превышения порога, ширина окна и другие, закладывается в процессе обучения.

Подсистема работы с нейронными сетями поддерживает загрузку уже существующих нейронных сетей для продолжения их обучения, что полезно при получении свежих данных из уже изученных сетей.

7. ЗАКЛЮЧЕНИЕ

В данной работе описана общая структура системы, предсказывающей превышение заданного порога значением длины трафика через заданный интервал времени, а также используемые данные и методы, использующиеся в их обработке.

Ниже приведены результаты исследований на тестовой выборке, извлеченной из набора данных, описанного во 2 части, которая состоит из 1700 временных шагов, следующих друг за другом.

Для получения следующих результатов, нейронная сеть, описанная в части 3 обучалась в течении 600 эпох с размером пакетов 2. Величина окна прогноза была выбрана 4. Обучающий набор содержал 1500 временных отсчетов, а тестовый – 200.

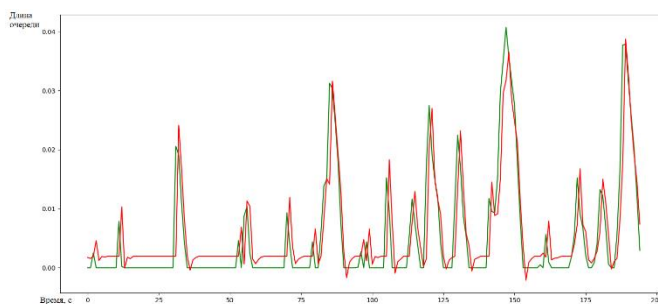


Рис. 1. Результат работы нейронной сети на тестовой выборке

Таким образом, было получено 196 тестовых прогнозов. Среди которых 11 ложно положительных и 4 ложно отрицательных прогнозов переполнения.

На рисунке (1) красной линией соединены значения длины очереди, предсказанные нейронной сетью, а зелёной – действительные значения, взятые из тестовой выборки. Значение по вертикальной оси нормализовано.

ЛИТЕРАТУРА

- [1] Paxson, V. Wide-Area Traffic: The Failure of Poisson Modeling / V. Paxson, S. Floyd // IEEE ACM Transactions on Networking. – 1995. – Vol. 3(3). – P. 226-244.
- [2] Leland, W.E. High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection / W. Leland, D.V. Wilson // Proc. IEEE INFOCOM '91. – Bal Harbor, FL, 1991. – P. 1360-1366.
- [3] Mondragón, R.J. Chaotic maps for traffic modelling and queueing performance analysis / R.J. Mondragón, D.K. Arrowsmith, J.M. Pitts // Performance Evaluation. – Netherlands, 2001. – P. 223-240.
- [4] Willinger, W. A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks / W. Willinger, M.S. Taqqu, A. Erramilli // Clarendon press. – Oxford, 1996. – P. 339-366.
- [5] CAIDA Data – Overview of Datasets, Monitors, and Reports [Электронный ресурс]. – Режим доступа: <https://www.caida.org> (20.01.2022).
- [6] Zhao, Z. LSTM network: a deep learning approach for short-term traffic forecast / Z. Zhao, W. Chen // IET Intelligent Transport Systems. – 2017. – Vol. 11(2). – P. 68-75.
- [7] Слой нормализации [Электронный ресурс]. – Режим доступа: <https://ru.keras.com/norm-layers/> (01.02.2022).
- [8] Libpcap File Format [Электронный ресурс]. – Режим доступа: <https://wiki.wireshark.org/Development/LibpcapFileFormat> (20.01.2022).