

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ МОНИТОРИНГА СОСТОЯНИЯ ОБОРУДОВАНИЯ ГЛОБАЛЬНЫХ СЕТЕЙ ЭВМ С ДИНАМИЧЕСКОЙ МАРШРУТИЗАЦИЕЙ

1. Введение

Современные средние и крупные территориально-распределенные сети ЭВМ ежесекундно формируют в среднем более 50 сообщений о неисправностях. Коррелировать сообщения (т.е. выделять сообщения, относящиеся к одной причине поломки, и определять эту причину) вручную при такой интенсивности потока становится практически невозможно. Динамическая маршрутизация современных сетей еще больше затрудняет эту задачу. Поэтому актуальной является разработка алгоритмов автоматической корреляции сообщений о неисправностях. В данной статье предложен метод корреляции, основанный на использовании нейронных сетей.

2. Постановка задачи

Задачей данной работы является разработка методов корреляции сообщений об ошибках в глобальных сетях ЭВМ на основе искусственных нейронных сетей (НС).

3. Модель глобальной сети ЭВМ

Объектом мониторинга является глобальная сеть ЭВМ (в дальнейшем просто "сеть"). Сеть представляет собой множество телекоммуникационных устройств, которые могут обмениваться информацией и средства коммуникации между ними. Эти телекоммуникационные устройства в дальнейшем будем называть сетевыми элементами (СЭ).

Информация о состоянии сети представлена сообщениями, которые СЭ посылают системе мониторинга состояния оборудования (СМСО). Не все сообщения свидетельствуют о неисправности сети. Основной практический интерес представляют сообщения о неисправностях соединений между СЭ. Выделим три уровня соединений между СЭ в соответствии с моделью OSI [1, 2]:

– физическое соединение (ФС) – физические средства коммуникации между двумя соседними СЭ (кабеля, оптоволокно и т.д.); соответствует физическому уровню OSI;

– логическое соединение (ЛС) – логический канал между двумя соседними СЭ, посредством которого коммутируются пакеты (VPC/VCC в ATM, DLC в FR, каналы, образованные метками MPLS и т.п. [3, 4, 5]); соответствует каналному уровню OSI;

– сервисное соединение (СС) – абстрактное соединение между двумя СЭ (не обязательно соседними), предоставляющее услуги по передаче информации между ними (VPN, VoIP, маршрутизация сетевых протоколов, например, IP и др.); соответствует сетевому и более высоким уровням OSI.

Оконечные точки (ОТ) физического, логического и сервисного соединений обозначим соответственно ОТФС, ОТЛС, ОТСС. Они уникально идентифицируются в пределах каждого СЭ, например:

– ОТФС – слот/порт или номер интерфейса низкого уровня;

– ОТЛС – VPI/VCI, DLCI или метки MPLS;

– ОТСС – VPN-порт или номер интерфейса высокого уровня.

Физической топологией (ФТ) назовем граф следующего вида. Узлами графа являются ОТФС и СЭ, причем каждая ОТФС связана с СЭ, которому она принадлежит. ОТФС разных СЭ также могут быть связаны между собой, причем каждое связующее ребро соответствует некоторому ФС. Сервисную топологию (СТ)

определим аналогично физической топологии (ФТ). Логическая топология (ЛТ) будет определена несколько иначе, чем ФТ и СТ. В реальной сети разводка кабелей (физический уровень) и предоставляемые клиентам сервисы (сервисный уровень) практически не меняются независимо от того исправны ФС и СС или нет. Естественно, с оговоркой, что устранить неисправность ФС и СС пытаются в кратчайшие сроки. ЛТ же нас интересует, в первую очередь, с точки зрения обеспечения маршрута между оконечными точками СС (такой маршрут, представляющий собой цепочку ЛС, назовем логическим туннелем). Поэтому ЛС будут считаться ребрами ЛТ только в том случае, если они в данный момент времени находятся в работоспособном состоянии и входят в некоторый логический туннель. В остальном определение ЛТ аналогично определениям ФТ и СТ. В результате такого определения ФТ и СТ сети слабо меняются со временем, зато ЛТ может существенно изменяться за счет создания/удаления ЛС при динамическом изменении маршрутизации. Принадлежность ЛС логическим туннелям также может сильно меняться при изменениях маршрутизации.

Таким образом, получается трехуровневая модель сети, физический уровень которой представлен ФТ, логический – ЛТ и сервисный – СТ. При этом каждый более низкий уровень предоставляет услуги по передаче информации более высокому. Так, СС реализуется при помощи логического туннеля. Каждое ЛС использует некоторое ФС для непосредственной передачи данных. При этом каждый логический туннель и каждое ЛС может быть одновременно использовано многими СС, а каждое СС, в принципе, может одновременно пользоваться несколькими логическими туннелями и ЛС, т.е. отношение между СС и ЛС типа многие-ко-многим. Каждое ФС может одновременно использоваться многими ЛС, но каждое ЛС в данный момент времени может использовать только одно ФС (хотя при изменении маршрутизации то же самое ЛС может начать использовать другое ФС), т.е. отношение между ЛС и ФС типа многие-к-одному. Модель сети показана на рис. 1 и рис. 2.

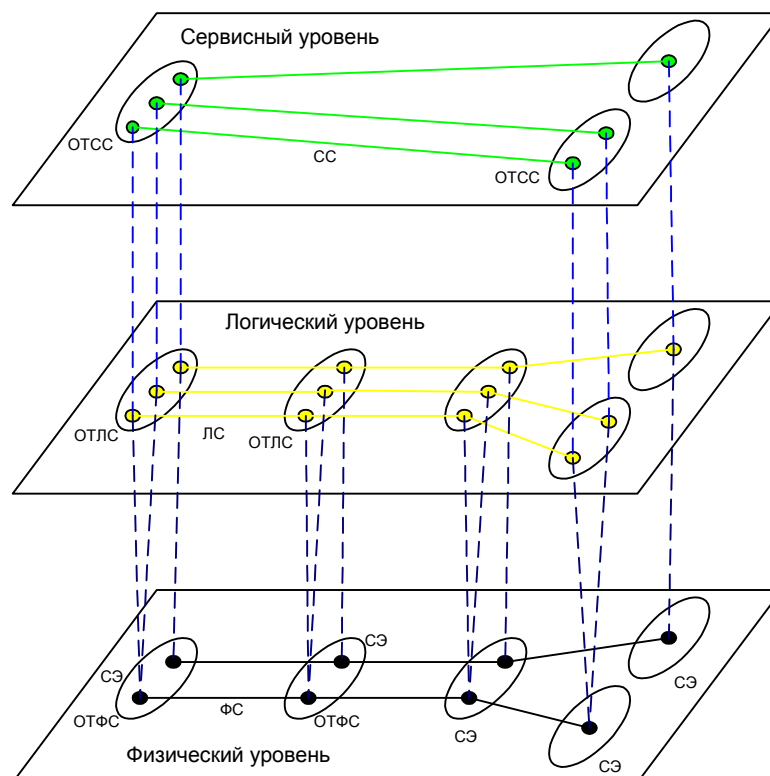


Рис. 1. Модель сети с выделением сетевых уровней

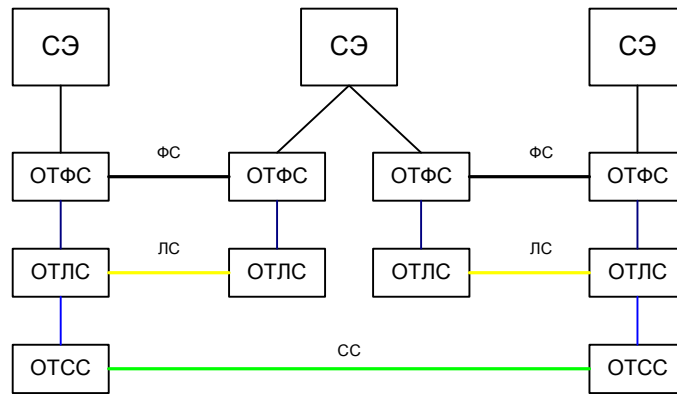


Рис. 2. Модель сети в виде графа

4. Модель приходящих в СМСО сообщений

Непосредственными объектами, на которые в СМСО приходят из сети сообщения, могут быть:

- СЭ;
- ОТФС;
- ОТЛС;
- ОТСС.

Поскольку интерес представляют сообщения о неисправностях соединений в сети, выполним классификацию приходящих сообщений следующим образом:

- неисправность ФС (НФС) – обрыв кабеля и т.п.;
- неисправность ЛС (НЛС) – невозможность коммутации между соседними СЭ, как правило, вследствие НФС;
- неисправность СС (НСС) – невозможность выполнения услуг по передаче информации между произвольными СЭ, как правило, вследствие НЛС (отсутствия связной цепочки ЛС между отправителем и получателем информации);
- другие сообщения (которые в данной модели во внимание не принимаются).

Соответствие между классами сообщений и объектами, на которые они приходят, приведено в таблице 1.

Таблица 1. Соответствие между классами сообщений и объектами неисправности

Класс сообщений	Объект неисправности
НФС	ОТФС
НЛС	ОТЛС
НСС	ОТСС

5. Модель генерирования сообщений о неисправностях

Наиболее распространенным случаем неисправности сети является НФС. Неисправность единственного физического соединения может вызвать появление большого количества сообщений о неисправности всех классов (НЛС, НСС, ИМ, ПК), так называемый "ураган ошибок".

Если же НФС отсутствует, то НЛС и НСС могут возникнуть при неправильных конфигурационных установках, что случается достаточно редко, легко локализуется и не вызывает большого количества сообщений о неисправности.

Рассмотрим теперь основные правила генерирования сообщений данной модели:

- при НФС генерируются НФС-сообщения на обе ОТФС этого ФС;

- при НЛС генерируются НЛС-сообщения на обе ОТЛС этого ЛС;
- при НСС генерируются НСС-сообщения на обе ОТСС этого СС;
- при НФС генерируются пары НЛС-сообщений (по одному для каждого ОТФС этого ФС) на все ОТЛС, которые использовали данное ФС;
- при НФС с удачной попыткой изменения маршрута НЛС-сообщения могут как посылаться (ATM), так и не посылаться (FR, MPLS).

При этом возможны случаи утери сообщений о неисправности в результате перегрузки каналов. В среднем теряется порядка 5% сообщений в сети.

6. Типичные сценарии

6.1. Неисправность физического соединения при отсутствии альтернативного маршрута

Рассмотрим конфигурацию сети, показанную на рис. 3.

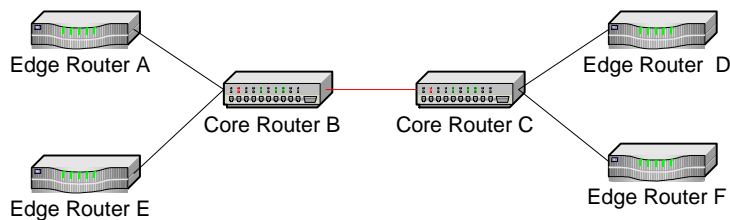


Рис. 3. Конфигурация сети с единственным маршрутом

Предположим, произошел отказ ФС между роутерами В и С. Тогда в сети будут сгенерированы следующие сообщения:

- по 1 НФС-сообщению на ОТФС роутеров В и С;
- по N НЛС-сообщений на ОТЛС роутеров В и С;
- по M_A , M_E , M_D и M_F НСС-сообщений на ОТСС роутеров А, Е, D и F соответственно (при отсутствии утери сообщений $M_A + M_E = M_D + M_F$).

6.2. Неисправность физического соединения при наличии альтернативного маршрута

Рассмотрим конфигурацию сети, показанную на рис. 4.

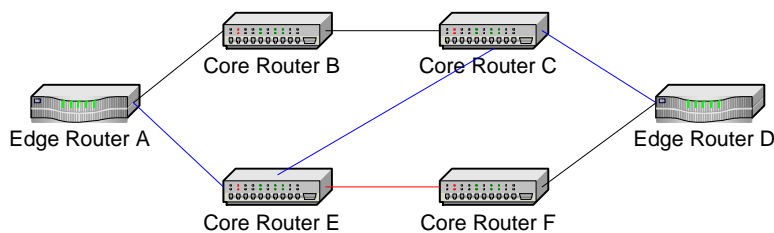


Рис. 4. Конфигурация сети с альтернативным маршрутом

В случае отказа ФС между роутерами Е и F – трафик, проходивший ранее через это соединение, будет динамически перемаршрутизирован. Предположим, будет выбран маршрут А-Е-С-Д. При этом будут сгенерированы следующие сообщения:

- по 1 НФС-сообщению на ОТФС роутеров Е и F;
- возможны два варианта в зависимости от технологии, используемой на канальном уровне:
 - по N НЛС-сообщений на ОТЛС роутеров В и С;
 - НЛС-сообщения сгенерированы не будут.

6.3. Отдельные неисправности логических соединений

Как было сказано выше, НЛС при отсутствии НФС – достаточно редкое явление. Сообщения, которые генерируются при НЛС, составляют часть сообщений, возникающих при НФС.

Неисправности логических соединений при отсутствии альтернативного маршрута

Рассмотрим конфигурацию сети, показанную на рис. 3. Если произошел отказ ЛС между роутерами В и С, то будут сгенерированы следующие сообщения:

- по 1 НЛС-сообщению на ОТЛС роутеров В и С;
- по M_A , M_E , M_D и M_F НСС - сообщений на ОТСС роутеров А, Е, D и F соответственно (без утери сообщений $M_A + M_E = M_D + M_F$).

Неисправности логических соединений при наличии альтернативного маршрута

Рассмотрим конфигурацию сети, показанную на рис. 4. Если произошел отказ ЛС между роутерами Е и F и маршрут был изменен на А-Е-С-D, то будут сгенерированы следующие сообщения:

- по 1 НЛС-сообщению на ОТЛС роутеров Е и F (в некоторых случаях такие сообщения вообще не генерируются).

7. Корреляция сообщений, выбор алгоритма установления корреляции

В задаче корреляции входящих из сети сообщений можно выделить следующие трудности [6]:

- шумы – из сети приходит много ненужных, избыточных, повторяющихся, случайных сообщений, происходят частые осцилляции (исправно-неисправно);
- скрытые зависимости – любая модель упрощает предметную область, поэтому появляются "невидимые" в данной модели элементы и связи, и их неисправность проявится в генерировании сообщений на другие элементы сети;
- сложные зависимости – реальные зависимости между элементами сети могут быть сложнее и разнообразнее, чем это представлено в модели;
- потеря сообщений;
- интенсивность потока сообщений: в сети может одновременно происходить большое количество неисправностей, что порождает поток сообщений высокой интенсивности (более 50 сообщ./сек.); это дополнительно усложняет корреляцию и накладывает строгие требования на производительность СМСО.

Одним из широко распространенных подходов к задаче корреляции является топологическая корреляция. В нем используется детализированная модель элементов сети и связей между ними. Основой построения алгоритма корреляции является так называемая база топологии – база данных, содержащая детальную информацию об элементах сети, их связях и зависимостях. Однако, в современных сетях применение топологической корреляции затруднительно в силу следующих причин:

- динамическая маршрутизация – автоматическое изменение маршрутизации при неисправностях в современных сетях приводит к быстрому устареванию базы топологии и, как следствие, к неверной корреляции;
- большое количество объектов сети – благодаря высокой степени детализации модели, в больших сетях база топологии может содержать несколько миллионов объектов и занимать порядка терабайта дискового пространства.

Применение нейронных сетей (НС) [7, 8] в задаче построения алгоритма корреляции сообщений имеет следующие достоинства:

– обучение на примерах – позволяет НС выявлять скрытые и сложные зависимости в сети, успешно бороться с шумами и неполнотой информации;

– обобщающие возможности – нет необходимости в детальной модели сети и знании полной информации об элементах сети и связях между ними (в том числе, в знании конкретной маршрутизации в данный момент времени), в процессе обучения НС имеет способность выделять существенные признаки для корреляции и абстрагироваться от малозначущих;

– компактное представление знаний – знания представлены в связях между нейронами в сжатой, компактной форме за счет абстрагирования от несущественных для корреляции признаков;

– эффективность – архитектура НС изначально ориентирована на параллельную обработку, что позволяет достичь высокой эффективности;

– дообучение – возможность дообучения в процессе эксплуатации для корректного отслеживания изменений в сети.

7.1. Тестовая телекоммуникационная сеть

Тестовая телекоммуникационная сеть состоит из 200 СЭ. Каждый СЭ имеет в среднем 20 слотов. Каждый слот включает в среднем 8 портов (ОТФС). У каждого порта – в среднем 20 ОТСС.

7.2. Предварительная обработка входящих сообщений

Основные задачи модуля предварительной обработки сообщений:

– фильтрация – в данной модели рассматриваются только НФС, НЛС и НСС-сообщения (см. раздел 4). Все остальные сообщения должны быть предварительно отфильтрованы, и задача их корреляции в данной работе не ставится;

– предварительная корреляция – все НЛС-сообщения коррелируются друг с другом и соответствующими ОТФС при помощи информации, доступной непосредственно из сообщения или путем запроса к СЭ, и не участвуют в дальнейшей НС-корреляции; принадлежность ОТФС к некоторому ФС задается базой физической топологии, таким образом, корреляция НФС-сообщений, отвечающих одному ФС, производится на основании базы физической топологии;

– компрессия – все одинаковые сообщения в пределах некоторого временного окна должны быть заменены на одно сообщение;

– пороговая фильтрация – фильтрация осциллирующих или случайных сообщений о неисправности, на которые в течение короткого промежутка времени приходит сообщение об устранении неисправности; при частом появлении подобных сообщений такое сообщение может быть передано в модуль НС-корреляции или может быть сгенерировано новое сообщение более высокого уровня;

– формирование входных векторов НС – входящие сообщения должны быть перекодированы в соответствии с форматом входов НС и поданы на соответствующие входы НС.

7.3. Предлагаемый нейросетевой подход

После предварительной обработки сообщений задача НС состоит в корреляции НСС-сообщений с соответствующими НФС-сообщениями. На основании критериев эффективности и времени обучения НС для целевой сети, параметры которой описаны в разделе 7.1, был выбран подход попарной корреляции сообщений.

Обобщающие возможности НС могут быть использованы благодаря следующим наблюдениям:

– все ОТСС граничного маршрутизатора, использующие один физический порт (или небольшую часть портов) маршрутизатора, имеют общую группу адресатов и, соответственно, общий набор альтернативных маршрутов через сеть;

– каждый потокообразующий коммутатор используется только некоторой частью СС сети, т.е. участвует в ограниченном наборе маршрутов.

На вход НС подается пара сообщений – НФС и НСС. При этом приходящие НФС-сообщения накапливаются в буфере, размер которого определяется размером временного окна. Размер временного окна определяется эмпирически и составляет около 900 сек. Все сообщения, абсолютная величина разности времен прихода которых превышает размер окна, считаются независимыми. Поэтому НФС-сообщения, разность времени прихода которых и текущего времени превышает по модулю размер временного окна, удаляются из буфера корреляции. Каждое приходящее НСС-сообщение коррелируется с каждым НФС-сообщением в буфере. В случае неуспешной корреляции оно коррелируется с каждым приходящим НФС-сообщением и считается независимым по истечении времени, равного размеру временного окна.

Поскольку НФС-сообщения, относящиеся к одному СС, коррелируются на этапе предварительной обработки сообщений, в корреляции и обучении НС участвует только одно из них. В случае утери оно должно быть сгенерировано в модуле предварительной обработки:

- на основании второго НФС-сообщения на данное СС;
- при превышении некоторого порога по количеству НФС-сообщений, использующих данное ОТФС.

7.3.1. Входы и выходы нейронной сети

Вход НС представлен двумя группами нейронов: одна – для НФС-сообщений, другая – для НСС-сообщений. На обе группы нейронов как для НФС, так и для НСС-сообщений подается идентификатор объекта поломки. Для НФС-сообщений – это физический порт (ОТФС), для НСС-сообщений – это ОТСС. Для кодирования объекта поломки использована информация об иерархической организации СЭ. Так, ОТФС идентифицируется номером СЭ, номером слота и номером порта. ОТСС, кроме этого, идентифицируется номером ОТСС в пределах данного СЭ. В соответствии с вышеприведенными наблюдениями, наиболее важная информация для НФС-сообщения – идентификатор СЭ, а для НСС-сообщения – номер физического порта данного СЭ. За правильную нумерацию СЭ отвечает модуль предварительной обработки сообщений. Все входные нейроны НС – бинарные, на них подается бинарный сигнал: +1 или –1. Кроме этого, на вход НС подается разность времен прихода НФС и НСС-сообщений, которая кодируется 20 бинарными входами (от –1000 сек. до +1000 сек. с точностью до трех знаков после запятой). Формат входов показан в таблице 2.

Таблица 2. Формат входов НС

Тип сообщения	Подчасть идентификатора объекта поломки	Количество нейронов
НФС	СЭ	24
	Слот	5
	Порт	3
НСС	СЭ	16
	Слот	5
	Порт	8
	ОТСС	5
Разность времен	Разность времен	20

Как видно из таблицы 2, кодирование СЭ осуществляется избыточно (чтобы закодировать 200 значений достаточно 8 бинарных входов), причем в случае НФС используется 24 входа, а в случае НСС – 16. Кодирование портов НСС также осуществлено 8 входами вместо возможных 3-х. Такой формат входов

позволяет улучшить обобщающие способности НС в соответствии с представленными в разделе 7.3 наблюдениями. Общее количество бинарных входов равно 86.

Выходом НС должна быть информация о том, является ли поданное на вход НСС-сообщение следствием данного НФС-сообщения. Однако, классификация на два класса затруднила бы обучение НС. Поэтому на выход подается идентификатор ОТФС, отвечающий данному НФС при успешной корреляции или – 1 на всех выходных нейронах, если сообщения независимы. Таким образом, НС обучается определять, используется ли данное ОТФС в возможных маршрутах сервисного соединения, отвечающего данному НСС. Для кодирования выхода НС используется 32 бинарных нейрона, упорядоченных аналогично нейронам, кодирующим входное НФС-сообщение.

7.3.2. Топология и обучающая выборка нейронной сети

Для реализации НС была выбрана топология многослойного персептрона с одним скрытым слоем как наиболее эффективная по быстродействию в рабочем режиме. Для обучения нейронной сети был выбран эффективный алгоритм семейства алгоритмов обратного распространения ошибки – Resilient Propagation. В качестве функции активации нейронов НС была выбрана сигмоида.

Обучающая выборка должна формироваться на истории сообщений о неисправностях данной сети. Оценка размеров выборки основывалась на следующих соображениях: каждое из $(200 \cdot 20 \cdot 8) / 2 = 16000$ ФС сети может использоваться 20-30 группами СС. Итого получается порядка $16000 \cdot 20 = 320000$ элементов в выборке. Чтобы предоставить НС информацию о независимых сообщениях, это число следует дополнить входными образами, отвечающими независимым НФС и НСС-сообщениям. Общее число таких "зависимых" и "независимых" входных образов следует сделать примерно равным. Таким образом, размер обучающей выборки составит порядка 600000 образов.

Выбор количества нейронов в скрытом слое НС осуществляется, как правило, эмпирически и на основе интуитивных соображений. Оценить размер скрытого слоя можно следующим образом. Количество синаптических весов в НС должно быть в 5-10 раз меньше, чем количество входных образов, т.е. порядка 100000. Разделив эту величину на суммарный размер входного и выходного слоев ($86 + 32 = 118$), получим скрытый слой примерно в 1000 нейронов. Эта величина будет использована в дальнейшей оценке эффективности НС.

7.3.3. Оценка эффективности нейронной сети

НС с конфигурацией в 86 входных нейронов, 1000 нейронов в скрытом слое и 32 выходными нейронами обрабатывает около 100 входных образов в секунду на персональном компьютере на базе Pentium II 400 МГц, 256 МВ RAM. Средней мощности серверная платформа позволит ускорить обработку в 10-20 раз. Найдем интенсивность потока сообщений, с которым будет справляться НС.

Обозначим через $I_{\text{ФС}}$ интенсивность потока НФС-сообщений, через $I_{\text{СС}}$ – интенсивность потока НСС-сообщений, через T – размер временного окна (около 900 сек.) и через $N_{\text{ФС}}$ – среднее количество НФС-сообщений в буфере. Они связаны следующими простыми соотношениями:

$$I_{\text{СС}} = C \cdot I_{\text{ФС}}; \quad (1)$$

$$N_{\text{ФС}} = I_{\text{ФС}} \cdot T, \quad (2)$$

где константа C определяет среднее количество вышедших из строя СС в результате неисправности ФС и примерно равна 40.

Пришедшее НСС-сообщение сравнивается с каждым НФС-сообщением в буфере, пока не будет скоррелировано. Среднее количество таких сравнений I в секунду можно оценить, как

$$I = \frac{I_{CC} \cdot N_{\phi C}}{2} . \quad (3)$$

Воспользовавшись (1) и (2), грубая оценка количества обрабатываемых НС входных образов в секунду:

$$I = \frac{C \cdot I_{\phi C}^2 \cdot T}{2} . \quad (4)$$

Быстродействие НС $I_{НС}$ должно превышать эту величину. Принимая $I_{НС} = 1000$ образов в секунду, получим:

$$I_{\phi C} < \sqrt{\frac{2 \cdot I_{НС}}{C \cdot T}} . \quad (5)$$

Подставляя значения, получаем: $I_{\phi C} < 0.25$ сообщений в секунду, но так как половина НФС-сообщений будет скоррелирована при предварительной корреляции, результатом будет пропускная способность коррелятора в 0.5 сообщений в секунду.

С учетом НПС и НСС-сообщений поток сообщений о неисправностях, эффективно обрабатываемых коррелятором, более, чем в 40 раз интенсивнее, то есть составляет около 20-30 сообщений в секунду.

8. Заключение

В данной статье предложен нейросетевой подход к корреляции сообщений о неисправностях в глобальной сети ЭВМ с динамической маршрутизацией. Описаны модель глобальной сети ЭВМ, модель сообщений о неисправностях в глобальной сети ЭВМ, модель генерирования сообщений о неисправностях и типичные сценарии неисправности. Показаны преимущества нейросетевого подхода к корреляции сообщений о неисправностях по сравнению с классическим топологическим методом, а именно: обучение на примерах, обобщающие возможности, компактное представление знаний, эффективность и возможность дообучения нейронной сети. Приведены оценки сложности архитектуры и размеров обучающей выборки нейронной сети. Поток эффективно обрабатываемых нейронной сетью сообщений о неисправностях оценивается в 20-30 сообщений в секунду. Такой эффективности достаточно для мониторинга даже относительно крупных сетей ЭВМ.

СПИСОК ЛИТЕРАТУРЫ

1. Information technology. Open Systems Interconnection. Basic reference model: The basic model. – ITU-T Recommendation X.200. – 1994. – P.1 – 63.
2. Management framework for Open Systems Interconnection (OSI) for CCITT applications. – ITU-T Recommendation X.700. – 1992. – P.1 – 11.
3. Микуцкий А. Основы технологии ATM // http://www.citforum.ru/nets/articles/atm_base.shtml.
4. Основы Frame Relay // <http://www2.bilim.com/1251/library/FR-basics.shtml>.
5. Always V. Advanced MPLS Design and Implementation. – Cisco Press, 2002. – Chapter 3.
6. Meira D.M. A Model for Alarm Correlation in Telecommunications Networks. – Belo Horizonte, 1997. – P. 1 – 169.
7. Уоссермен Ф. Нейрокомпьютерная техника. – М.: Мир, 1992. – С.1 – 184.
8. Fine T. Feedforward Neural Network Methodology. – Springer, 1999. – P. 1 – 340.