

РАЗДЕЛ V. СОВЕРШЕНСТВОВАНИЕ ИНСТРУМЕНТАРИЯ НА БАЗЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Нестерук Л.Г.

ПОВЫШЕНИЕ ОПЕРАТИВНОСТИ ИНКРЕМЕНТНЫХ КЛАССИФИКАТОРОВ В СОСТАВЕ АДАПТИВНЫХ СРЕДСТВ ЗАЩИТЫ ИНФОРМАЦИИ

(СПбГУЭФ, Санкт-Петербург)

Рассмотрен подход к организации нейросетевых классификаторов с инкрементным обучением, позволяющий повысить производительность интеллектуальных средств классификации в составе системы адаптивной защиты информации за счет минимизации количества кластеров в процессе периодического выполнения этапов структурной и параметрической оптимизации.

Атуальность. Инкрементное обучение позволяет реализовать обучение средств нейросетевой классификации в режиме реального времени, что принципиально важно для систем защиты информации (СЗИ), предназначенных для обеспечения безопасности критически важных объектов. К критически важным объектам информационно-телекоммуникационных структур, финансовой и банковской сферы, гидротехнических сооружений, нефтегазовых и химических вредных производств, АЭС и т.п. относят техногенные системы, выведение из строя которых может привести к катастрофическим последствиям. Одним из основных требований к СЗИ для критически важных объектов является оперативность реакции на угрозы деструктивных системных событий. В основе нейросетевых классификаторов с инкрементным обучением лежит биологический *принцип двойной пластичности*: структурные изменения в биосистеме происходят реже, чем изменения функциональных параметров, который, в первую очередь, реализуется чрез адаптацию функциональных параметров классификатора (например, значений весов межнейронных связей), а во вторую очередь – посредством модификации его структуры. Проблема классификации рассматривается как задача адаптации в оперативном режиме согласно принципу двойной пластичности, т.к. одновременно формируется топология и производится обучение нейронной сети (НС). К примеру, основная идея нейросетевого классификатора EMANN – первичность эволюции внутренних функциональных параметров классификатора в процессе обучения и вторичность изменения ее структуры. Улучшение точности классификатора сопровождается стремлением к поддержанию минимальной топологии НС.

Классификатор строится добавлением и/или изъятием формального нейрона (ФН) без существенного изменения весов существующих в сети межнейронных связей (информационного поля НС). Причем изменение структуры классификатора не сопровождается переобучением всего информационного поля НС, а только его отдельного фрагмента, связанного с упомянутым ФН, что увеличивает скорость выполнения операции классификации. Инкрементным

обучением с учителем при решении классификационных задач характеризуется класс нейронных сетей ART: сеть ARTMAP позволяет классифицировать входные данные по набору признаков в виде n -мерного вектора бинарных значений, представляющих наличие или отсутствие каждого признака; сеть Fuzzy ARTMAP (FAM) классифицирует входные данные по нечеткому набору признаков – вектору, элементами которого являются значения в диапазоне от 0 до 1, представляющие степень достоверности того или иного признака.

Основными проблемами сетей семейства ART являются монотонное возрастание числа кластеров в процессе эксплуатации классификатора, а также явление пролиферации, которое приводит к формированию большого числа практически идентичных кластеров, соответствующих близким входным векторам. Рост числа кластеров снижает оперативность выполнения операций классификации и обучения сетей ART, а следовательно, приводит к увеличению времени реакции СЗИ на динамику угроз.

Цель публикации – рассмотрение подхода к организации адаптивного инкрементного классификатора, в котором количество кластеров в процессе его эксплуатации будет увеличиваться не монотонно, т.е. предусматривается удаление избыточных формальных нейронов. Согласно предложенному подходу на этапе структурной пластичности помимо увеличения числа кластеров в процессе обучения НС будет периодически выполняться процедура сокращения их числа за счет удаления малозначачих для классификации и идентичных по параметрам ФН (кластеров, сформированных посредством пролиферации). Остановимся на архитектуре сетей FAM и EMANN.

Fuzzy ARTMAP часто реализуют в виде упрощенной модели (рис. 1), полученной путем комбинации самообучаемой сети ART с полем преобразования (map field).

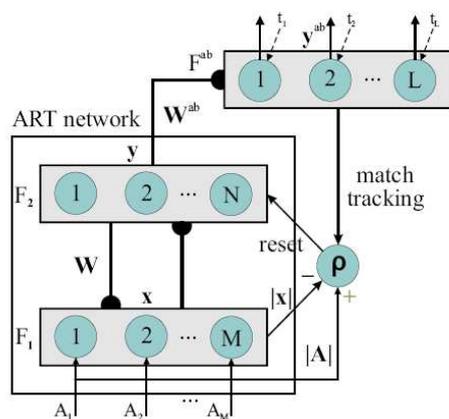


Рис. 1. Архитектура Fuzzy ARTMAP

FAM состоит из двух уровней узлов (нейронов) с полными связями: M узлов входного слоя F_1 , и N узлов соревновательного слоя F_2 . Набор вещественных весов.

Нейронная сеть EMANN. Архитектура НС представляется в виде иерархии модулей и подобно методу каскадных корреляций вводится операция расширения НС (рис. 2).

Каждый модуль EMANN – НС с одним скрытым слоем. Входы EMANN соединены с входами всех вышележащих модулей, что позволяет избежать зависимости новых модулей от имеющихся модулей. Выходы имеющихся модулей соединены с входами вновь введенного модуля, а решение задачи снимается с выходов модуля верхнего уровня. В процессе обучения (параметрическая пластичность) каждый новый модуль должен дополнять существующие и извлекать полезную информацию из модулей более низкого уровня.

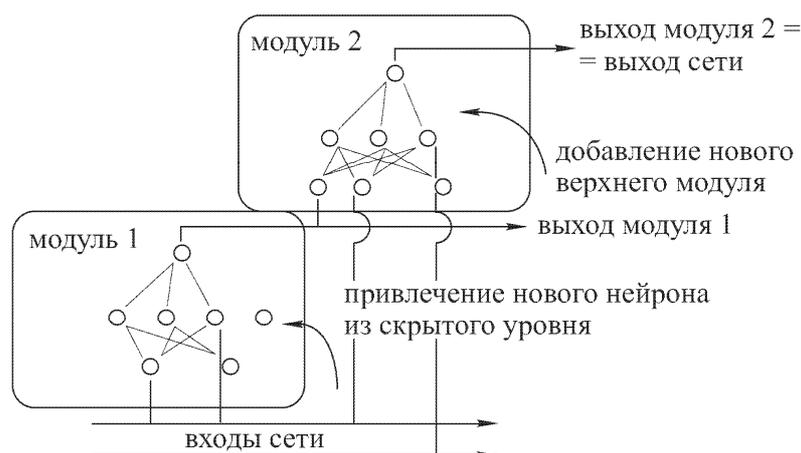


Рис. 2. Модульная архитектура EMANN

Минимальная структура EMANN имеет один ФН в скрытом слое. Количество входов и выходов определяется условиями задачи. Так как начальная структура (информационное поле НС) минимальна, то для поддержания обобщающей способности следует периодически наращивать информационное поле НС путем добавления ФН и соответствующих межнейронных связей (структурная пластичность). Также периодически следует избавляться от малозначачих для классификации ФН [3]. Для оптимизации весов, т.е. параметрической пластичности, в EMANN используется алгоритм extended delta-bar-delta – модификация метода Back Propagation.

Широкое использование сетей семейства ART позволяет считать актуальным поиск подходов, позволяющих улучшить их эксплуатационные характеристики и, не в последнюю очередь, скорость выполнения операций классификации и инкрементного обучения. Реализация в инкрементных классификаторах семейства ART процедуры сокращения малозначачих и повторяющихся кластеров позволяет увеличить производительность построенных а их основе адаптивных средств защиты информации как в режиме работы, так и в процессе обучения, а, следовательно, в полной мере реализовать режим функционирования СЗИ, близкий к реальному масштабу времени.

Щадилов А.Е., Исаев В.И., Швецов Е.А.

СОЦИАЛЬНЫЕ СЕТИ КАК БИЗНЕС-СРЕДА

(СПбГУЭФ, Санкт-Петербург)

Социальная сеть – это интерактивный многопользовательский веб-сайт, контент которого наполняется самими участниками сети. Социальные сети стремительно набирают популярность во всём мире, аудитории некоторых из них превышают население целых стран. В настоящее время сложно найти чело-