УДК 004.032.26:004.052

И.В. Потапов

ИССЛЕДОВАНИЕ И ОПТИМИЗАЦИЯ НАДЕЖНОСТИ НЕЙРОКОМПЬЮТЕРНЫХ СИСТЕМ В КОНФЛИКТНЫХ СИТУАЦИЯХ

Рассматриваются задачи исследования и оптимизации надежности отказоустойчивых нейрокомпьютерных систем на базе избыточных искусственных нейронных сетей. Постановки задач сформулированы в терминах теории игр. Решения рассматриваемых задач ориентированы на использование ПЭВМ.

Ключевые слова: *Избыточные нейронные сети*, *оптимизация надежности*, *игровое моделирование.*

Учитывая специфические особенности решаемых нейрокомпьютерами задач (как правило, это трудно формализуемые задачи, эффективное решение которых с помощью традиционных средств вычислительной техники затруднено), а также существенный выигрыш во времени при решении задач обработки информации за счет характерного для искусственных нейронных сетей (ИНС) массового параллелизма вычислений, можно полагать, что в ближайшем будущем нейровычислительные устройства будут широко применяться в области так называемых технологий двойного назначения. В этом случае должны предъявляться повышенные требования к функциональной надежности ИНС, являющихся основным вычислительным компонентом нейрокомпьютерных систем (НКС). Решение задач противоборства НКС имеет важную фундаментальную и прикладную значимость, поскольку в системах двойного назначения не исключены конфликтные ситуации, в которых ИНС нейрокомпьютерных систем могут целенаправленно подвергаться атакам, преследующим цель дестабилизировать работу нейрокомпьютеров. Исследование надежности функционирования ИНС в конфликтных ситуациях удобно проводить с использованием аппарата математической теории игр.

В данной работе объектом исследования является многослойная многовыходная избыточная структурно-однородная адаптивная к отказам «стареющая» искусственная нейронная сеть $S_{\mathrm{A}}(n,m,\mathbf{s})$ [1], допускающая динамическое перераспределение однородных универсальных резервных блоков искусственных нейронов (ИН) между q группами нейронных блоков. В момент включения резервного нейронного блока вместо основного система контроля и адаптации НКС выполняет его настройку, после чего подключенный резервный блок ИН начинает работать в режиме замененного основного блока и подвергаться воздействию того же потока отказов, что и остальные элементы соответствующей группы.

2. Задача оптимизации резервирования структурно-однородной ИНС

Пусть ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$ вступает в противоборство с противником, преследующим цель дестабилизировать ее работу. В качестве средств атаки противник имеет возможность увеличивать интенсивности отказов $\boldsymbol{\lambda}(t)=(\lambda_i(t))$ в группах ИН атакуемой ИНС и интенсивность отказов не включенных в работу резервных

блоков $\lambda_0(t)$, т.е. усиливать «старение» блоков ИН. Для защиты от атаки система контроля и адаптации (которую для простоты будем считать абсолютно надежной) может перераспределять имеющийся резерв между q атакуемыми группами ИН, т.е. формировать вектор резервирования $\mathbf{s}(t)$. Будем полагать, что ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$ может находиться в $(m+1)$ состояниях по числу k возможных отказов основных и резервных блоков ИН, причем состояния E_k ($0 \le k \le m$) являются работоспособными состояниями, а E_{m+1} – поглощающим состоянием. Тогда, учитывая нестационарные потоки отказов блоков ИН, вероятностная модель функционирования такой ИНС определяется системой дифференциальных уравнений

$$\mathbf{p}'(t) = D\mathbf{p}(t), \qquad (1)$$

где $\mathbf{p}(t)$ – вектор-столбец размерности $(m+1)$ вероятностей нахождения системы в работоспособных состояниях; D – матрица размерности $(m+1)\times(m+1)$ переменных коэффициентов дифференциальных уравнений вида [1], зависящих от вектора резервирования $\mathbf{s}(t)$, вектора интенсивностей отказов в группах $\boldsymbol{\lambda}(t)$ и распределения отказов в блоках ИН. В терминах теории оптимального управления $\mathbf{p}(t)$ является фазовым вектором, а $\mathbf{s}(t)$ и $\boldsymbol{\lambda}(t)$ – векторы управления.

Система (1) описывает дифференциальную игру, являющуюся моделью противоборства двух игроков, один из которых располагает стратегиями нападения $\boldsymbol{\lambda}(t)$, а второй – стратегиями защиты $\mathbf{s}(t)$. При этом на управления игроков (стратегии) в общем виде могут быть наложены естественные ограничения, определяемые физической сущностью и условиями функционирования системы. В качестве функции платы в игре рассматривается вероятность безотказного функционирования ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$, определяемая из решения системы (1).

Решением вышеописанной дифференциальной игры будут оптимальные управления игроков, позволяющие им получить максимальный выигрыш (минимальный проигрыш). Для машинного решения задачи оптимального динамического распределения резерва ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$ рассматриваемая задача путем дискретизации сводится к многошаговой матричной игре с соответствующими ограничениями на стратегии игроков [2]. Численное решение матричной игры находится в смешанных стратегиях методом многократного фиктивного разыгрывания на ПЭВМ.

3. Оптимизация надежности ИНС в условиях игры с «природой»

Вышеприведенная модель противоборства базировалась на том, что интересы сторон, участвующих в конфликтной ситуации, противоположны и атакующая сторона стремится причинить максимальный ущерб ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$, т.е. рассматривалась антагонистическая игра с разумным противником. Не менее важное фундаментальное и прикладное значение имеют так называемые игры с «природой», в которых противник («природа») не стремится причинить максимальный вред обороняющейся стороне, а действует случайным образом. Использование в игре с «природой» стратегий, определенных в предположении заинтересованности обеих сторон в получении максимального выигрыша (минимального проигрыша), неэффективно, поскольку такие стратегии соответствуют принятию решений в условиях крайнего пессимизма в оценке действий «природы».

Пусть сторона A располагает вышеописанной «стареющей» ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$. На интервале времени [0, t_f], где t_f – время окончания игры, введем вектор $\boldsymbol{\tau}=(\tau_0,\tau_1,\tau_2,...,\tau_L)$, $\tau_0=0$, $\tau_L<t_f$, элементы которого соответствуют моментам перераспределения резервных нейронных блоков ИНС между q группами. В дальнейшем вектор $\boldsymbol{\tau}$ будем называть вектором настройки ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$. Каждому моменту времени τ_k ($0\le k\le L$) поставим в соответствие вектор распределения резервных элементов $\mathbf{s}^{\tau_k}=(s_1^{\tau_k},s_2^{\tau_k},...,s_q^{\tau_k})$. Таким образом, игрок A располагает стратегиями $W^A=\{\boldsymbol{\tau},s(\boldsymbol{\tau})\}$, удовлетворяющими естественным ограничениям на минимальное время между двумя соседними настройками и на использование ресурсов защиты, т.е количество перераспределяемых в моменты настройки резервных элементов.

Пусть сторона B («природа») может находиться в одном из состояний $B=\{B_1,B_2,...,B_N\}$, множеству которых ставится в соответствие множество интенсивностей отказов ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$ $\lambda(t)=\{\boldsymbol{\lambda}^i(t)\}$, $1\le i\le N$, каждый элемент которого представляет собой совокупность интенсивностей отказов в q группах основных блоков ИН и не включенных в работу резервных блоков $\boldsymbol{\lambda}^i(t)=\left(\lambda_0^i(t),\lambda_1^i(t),...,\lambda_q^i(t)\right)$. Пусть для каждого состояния «природы» $B_i\in B$ заданы вероятности нахождения в этих состояниях $\mathbf{Q}(t)=\{Q_1(t),Q_2(t),...,Q_N(t)\}$. Тогда множество стратегий «природы» можно определить как $W^B=\{\mathbf{Q}(t),\lambda(t)\}$ с соответствующим ограничениями, в том числе на суммарное нападение на нейронные блоки ИНС. В рассматриваемой игре действия «природы» заключаются в случайном выборе одного из N состояний, которым соответствуют интенсивности потоков отказов основных и резервных нейронных блоков ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$.

В качестве функции платы в рассматриваемой игре используется вероятность безотказной работы ИНС $S_{\mathrm{A}}(n,m,\mathbf{s})$ к моменту окончания игры $P(t_f)$. Тогда решением игры будут вектор $\boldsymbol{\tau}^{\max}$ моментов настроек ИНС и множество векторов резервирования $s^{\max}=\{\mathbf{s}(\tau_k)\}$, соответствующих моментам настройки τ_k ($0\le k\le L$), максимизирующие вероятность безотказной работы $P(t_f)$ рассматриваемой ИНС.

В общем случае такая игра описывается дифференциальной моделью, однако учитывая ограничения на управления игроков, в том числе особенности технической реализации распределения резервных элементов ИНС, дифференциальная игра сводится к многошаговой матричной игре, которую удобно решать с использованием ПЭВМ [3].

4. Оптимизация надежности в задаче противоборства НКС

В развитие рассмотренных выше моделей конфликтных ситуаций можно предположить, что в особых условиях нейрокомпьютерные системы (как разновидности интеллектуальных систем в области технологий двойного применения) могут оказаться способными вступать в конфликты и противоборство не только с «природой», но и между собой и активно влиять на работоспособность противо-

борствующей стороны, например, путем целенаправленного воздействия, приводящего к росту интенсивности отказов компонентов (блоков искусственных нейронов) нейрокомпьютерной системы, основу которой составляет искусственная нейронная сеть, т.е. в конечном итоге искусственно ускорять процесс «старения» нейронной сети этой системы в своих интересах. Такое взаимодействие НКС может быть осуществлено, например, путем взаимного вмешательства в процессы обучения и текущего дообучения нейронных сетей в процессе функционировании НКС, а также путем опосредованного влияния на физические условия функционирования противоборствующей стороны. Не останавливаясь подробно на способах взаимодействия НКС, рассмотрим фундаментальную задачу противоборства.

В качестве моделей НКС, участвующих в игре, будем рассматривать однородные избыточные восстанавливаемые ИНС $S_{\text{AB}}(n,m,\mathbf{s})$ [1,4]. Пусть игрок 1 располагает НКС $S^1_{\text{AB}}\left(n^1,m^1,\mathbf{s}^1\right)$, а игрок 2 располагает НКС $S^2_{\text{AB}}\left(n^2,m^2,\mathbf{s}^2\right)$. Обе нейрокомпьютерные системы $S^g_{\text{AB}}\left(n^g,m^g,\mathbf{s}^g\right)$, $g=1,2$, являются восстанавливаемыми с интенсивностями восстановления $\mu^g(t)$, одинаковыми для всех нейронных блоков. Под восстановлением может пониматься как логическая перестройка параметров ИН, так и выполнение процедур дообучения в условиях искажения или зашумленности выходных сигналов отдельных ИН блока. Игрок 1 располагает множеством стратегий $W^1=\left\{\mathbf{s}^1,\boldsymbol{\lambda}^2,\mu^1\right\}$, а игрок 2 располагает множеством стратегий $W^2=\left\{\mathbf{s}^2,\boldsymbol{\lambda}^1,\mu^2\right\}$, где $\mathbf{s}^g$ – вектор резервирования g-го игрока; $\boldsymbol{\lambda}^g=\left(\lambda^g_0(t),\ \lambda^g_1(t),...,\lambda^g_q(t)\right)$ – вектор интенсивностей отказов в группах нейронных блоков НКС $S^g_{\text{AB}}\left(n^g,m^g,\mathbf{s}^g\right)$.

Будем считать, что число групп нейронных блоков ИНС для игроков 1 и 2 одинаково. Положим, что за время игры t_f игрок g $(g=1,2)$ имеет право не более чем $(L\geq 1)$ раз, не считая момента $t=0$, изменять вектор резервирования $\mathbf{s}^g$ управляемой им системы $S^g_{\text{AB}}\left(n^g,m^g,\mathbf{s}^g\right)$. Вектор $\boldsymbol{\tau}^g=\left(\tau^g_0,\tau^g_1,...,\tau^g_L\right)$, $\tau^g_0=0$, $\tau^g_L<t_f$, назовем вектором настройки g-го игрока. Будем полагать, что в процессе противоборства игрок 1 стремится максимизировать величину $P^1(t_f)-P^2(t_f)$, являющуюся функцией платы, где t_f – время окончания игры, $P^g(t)$ – вероятность безотказной работы g-й ИНС, а игрок 2 стремится минимизировать эту величину.

Для исследования задачи противоборства двух нейрокомпьютерных систем $S^g_{\text{AB}}\left(n^g,m^g,\mathbf{s}^g\right)$, $g=1,2$, необходимо решить игру двух лиц с нулевой суммой, где функция выигрыша $G(z_1,z_2)=P^1(t_f)-P^2(t_f)$. В приведенных обозначениях $z_g\in W^g$ – стратегия g-го игрока.

Дифференциальная игра двух нейрокомпьютерных систем с учетом всех ограничений на ресурсы нападения (интенсивности отказов нейронных блоков ИНС противника) и защиты (интенсивности восстановления отказавших нейронных

блоков и количества резервных блоков ИН), минимального времени между двумя последовательными настройками систем и заданной точности изменения управлений игроков методом дискретизации сводится к матричной игре заданной размерности, имеющей нормальную форму [1,4]. Машинное решение игры ищется в смешанных стратегиях методом многократного разыгрывания.

Заключение

В рассмотренных выше постановках задач оптимизации надежности нейровычислительных систем в качестве оптимизируемого функционала выбрана одна из важнейших характеристик надежности – вероятность безотказной работы системы. На практике возможны ситуации, в которых в качестве оптимизируемого функционала необходимо использовать другие характеристики надежности, которые не всегда могут быть получены непосредственно из решения системы (1). Так, например, оптимизация надежности по среднему времени работы системы до отказа (среднему времени «жизни» системы) потребует усложнения процедуры расчета функции платы, поскольку, в отличие от задач статической оптимизации надежности, в рассматриваемых задачах данный параметр должен рассчитываться по-другому, например, как момент времени, для которого математическое ожидание числа отказавших нейронных блоков системы равно $(m+1)$, где m – число резервных блоков.

ЛИТЕРАТУРА

1. *Потапов И.В.* Надежность нейрокомпьютерных систем. Модели и задачи. Омск: Изд-во ОмГТУ, 2007.
2. *Потапов И.В.* Решение задачи оптимального динамического распределения резерва «стареющей» искусственной нейронной сети в конфликтной ситуации // Нейрокомпьютеры: разработка, применение. 2006. № 3. С. 3 – 8.
3. *Потапов И.В.* Резервирование «стареющей» искусственной нейронной сети в условиях игры с «природой» // Надежность. 2006. № 4(19). С. 3 – 10.
4. *Потапов В.И.*, *Потапов И.В.* Противоборство (дифференциальная игра) двух нейрокомпьютерных систем // Информационные технологии. 2005. № 8. С. 53 – 57.

Статья представлена кафедрой программирования факультета прикладной математики и кибернетики Томского государственного университета и оргкомитетом 7-й Российской конференцией с международным участием «Новые информационные технологии в исследовании сложных структур», поступила в научную редакцию 9 октября 2008 г.