

Воробьева А.А., Крючин О.В.
Тамбовский Государственный Университет им. Г.Р. Державина
Россия, г. Тамбов
tmiooy@gmail.com

Исследование оптимального числа элементов информационных ресурсов для построения моделей, базирующихся на технологии искусственных нейронных сетей

В настоящее время использования большого числа информационных ресурсов для обучения искусственных нейронных сетей (ИНС) становится все более актуальным. Это связано с одной стороны с развитием технологий, значительно упрощающих использование ИНС, а с другой — со значительными (часто недопустимо большими) временными затратами на построение моделей. В качестве информационных ресурсов могут выступать вычислительные кластерные или распределенные системы. Как показано в работах [1-3], алгоритм обучения ИНС позволяет произвести распределение нагрузки на четырех основных уровнях, а именно:

1. при вычислении значения невязки [1];
2. при подборе весовых коэффициентов (алгоритм в значительной степени зависит от выбранного метода) [3];
3. при подборе активационных функций нейронов [2];
4. при подборе числа скрытых слоев сети [1].

Важность интерконнекта между элементами информационных ресурсов (ИР-элементов), то есть между узлами кластерной системы или компьютерами, объединенными в распределенную систему, напрямую зависит от выбранного уровня. То есть, если все ИР-элементы задействованы для вычисления значения невязки, то обмен данными между ними будет весьма интенсивен. Напротив, при обучении каждым ИР-элементом своей структуры сети передача данных сводится к минимуму.

Как известно, эффективность информационных процессов обучения ИНС при использовании большого числа ИР-элементов можно оценить двумя различными способами — эмпирическим (путем измерения временных затрат) и аналитическим (вычисление числа мультипликативных операций, которые необходимы для обучения). В обоих случаях коэффициент эффективности прямо пропорционален отношению числа временных затрат либо количеству операций для последовательно обучения к произведению количества используемых ИР-элементов на число затрат для распределенного (или параллельного) обучения.

$$\alpha_t = \frac{t}{n\tau(n)} \quad (1)$$

$$\alpha_z = \frac{z}{nZ(n)} \quad (2)$$

В данных формулах t — время, затраченное последовательным информационным процессом, $\tau(n)$ — время, затраченное распределенным процессом при использовании n ИР-элементов, z — число операций, совершенных последовательным информационным процессом, $Z(n)$ — наибольшее число мультипликативных операций, которые выполняются на используемых ИР-элементах.

Разумеется, с расширением информационного ресурса (то есть с увеличением числа ИР-элементов) эффективность распределенных информационных процессов начинает несколько снижаться. В свою очередь это приводит к тому, что временные затраты начинают снижаться менее быстро. Несколько неожиданным является тот факт, что после определенного момента они начинают возрастать, поскольку при большом числе используемых ИР-элементов время, которое расходуется на передачу данных между ними, а также на синхронизацию ИР-элементов оказывается больше того времени, которое затрачивается на обучение каждым таким ИР-элементом. Это проиллюстрировано на рис. 1 и 2, где показаны зависимости временных затрат и значения коэффициента эффективности от числа используемых ИР-элементов.

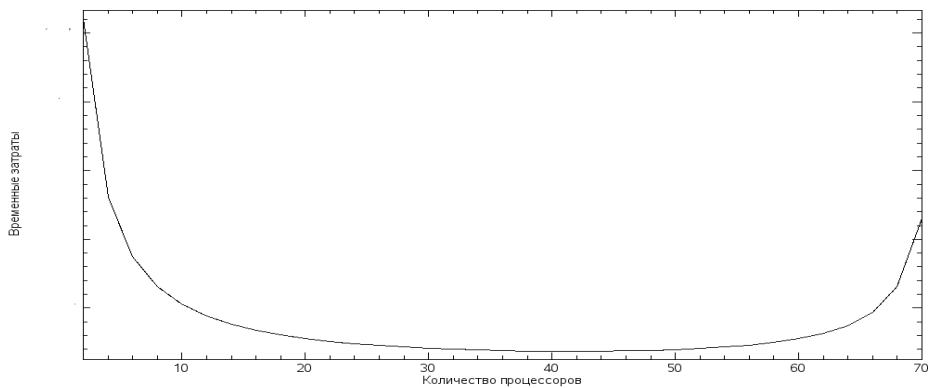


Рис. 1. График зависимости временных затрат от числа используемых ИР-элементов.

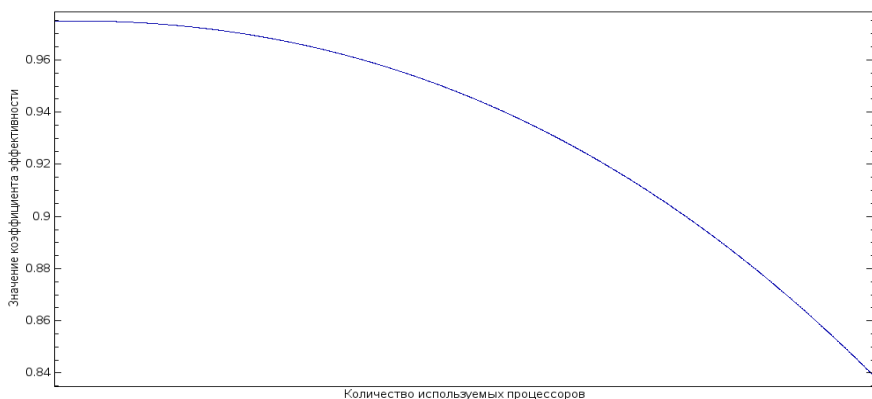


Рис. 2. График зависимости коэффициента эффективности от числа используемых ИР-элементов.

Использование распределенных и кластерных систем сопряжено с определенными затратами, размер которых часто пропорционален числу использованных ИР-элементов, то есть имеет место зависимость $\Theta_T = \Theta_T(n)$. В простейшем случае, когда считается, что стоимость использования ИР - элементов определяется лишь временем эксплуатации и числом задействованных ИР-элементов, эта зависимость может быть выражена как $\Theta_T \approx n\tau(n)$. Известно, что временные затраты $\tau(n)$ напрямую зависят от числа используемых ИР-элементов [4]. Исходя из этого зависимость стоимости использования ИР (такого как вычислительный кластер или распределенная система) от числа использованных ИР-элементов изображена на рис. 3.

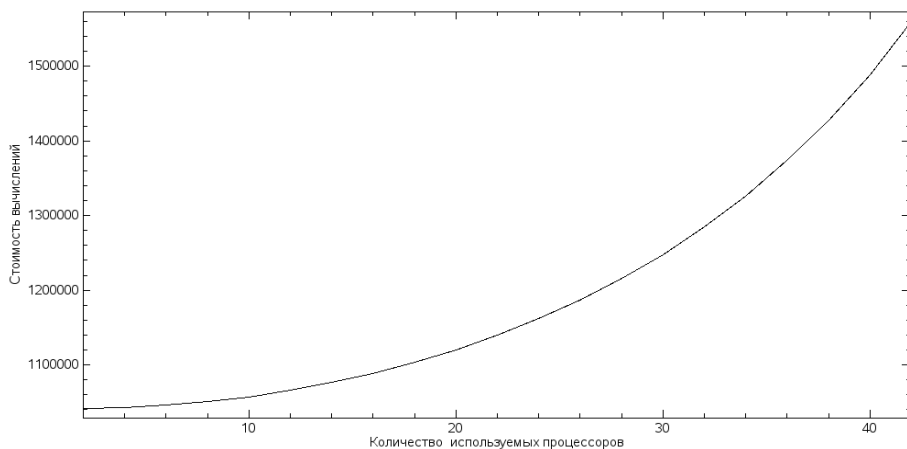


Рис. 3. График зависимости стоимости использования ИР от числа используемых ИР-элементов.

Однако, следует учитывать, что стоимость ожидания результатов также требует определен-

ных затрат и, следовательно, также зависит от числа используемых ИР-элементов $\Theta_U = \Theta_U(n)$.

Следовательно, затраты на ожидание результатов обучения ИНС прямо пропорциональных затраченному времени и, как следствие, снижаются с увеличением числа используемых ИР-элементов, но лишь до определенного предела (который можно условно обозначить как n_L) после чего вновь начинают возрастать. Это проиллюстрировано на рис. 4

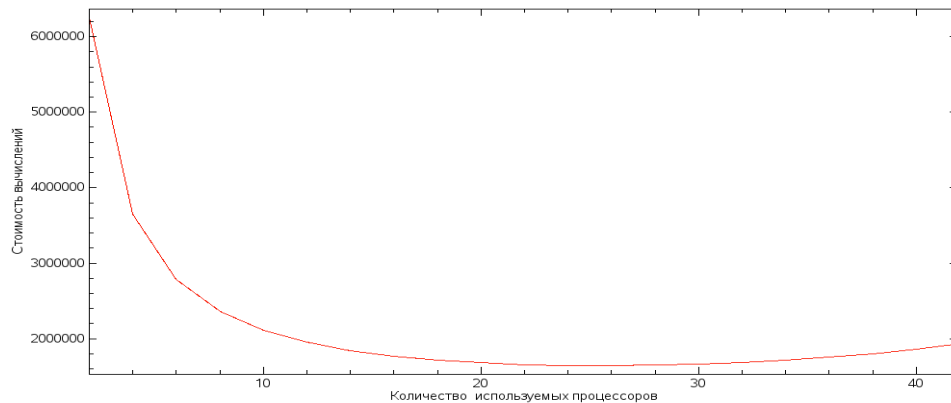


Рис. 4. График зависимости стоимости построения ИНС-модели от числа используемых ИР-элементов.

Таким образом, общие затраты на выполнение задачи построения ИНС-модели при использовании n ИР-элементов определяется формулой

$$\Theta_F = \Theta_T(n) + \Theta_U(n) \approx \theta n \tau(n) + q_t c_t + \tau(n) q_t \quad (3)$$

θ — стоимость использования 1 вычислительного узла в течение 1 часа, $\tau(n)$ — стоимость использования, c_t — время, необходимое на подготовку, q_t — стоимость ожидания 1 часа.

Исходя из вышесказанного, можно сделать вывод, что для любой задачи существует оптимальное число ИР-элементов, которые следует использовать для ее решения. Конкретное число ИР-элементов определяется особенностями моделируемого объекта, типом выбранной ИНС, а также информационными процессами распределенного обучения.

1. Литература

2. Крючин О.В. Построение модели информационных процессов выбора структур искусственной нейронной сети // Вестник ТГУ, т.19, вып.6, 2014 — С. 1836-1840.
3. Арзамасцев А.А., Крючин О.В. Реализация параллельного алгоритма подбора активационных функций искусственной нейронной сети // ЭФТЖ, т.6, 2011— С. 52-61.
4. Арзамасцев А.А., Крючин О.В., Слетков Д.В., Зенкова Н.А. Количество мультипликативных операций при подборе весовых коэффициентов искусственной нейронной сети // Вестник ТГУ, т.18, вып.1, 2013 — С. 176-182.
5. Крючин О.В. Информационные процессы обучения искусственных нейронных сетей методом полного сканирования и аналитическая модель параллельной версии этих процессов // Вестник ТГУ. Серия Естеств. и техн. Науки.-Тамбов, 2015— С. 142-145.