

УПРАВЛЕНИЕ ВЕЛИЧИНОЙ ОШИБКИ В НЕЙРОННЫХ СЕТЯХ

© 2016 А.В. Игнатенков¹, А.М. Ольшанский²¹ООО «Научно-технологический центр по эксплуатации железных дорог», г. Москва²Самарский государственный университет путей сообщения

Статья поступила в редакцию 11.11.2016

В статье представлены постановка и решение задачи управления ошибкой искусственной нейронной сети с переменной проводимостью сигнала.

Ключевые слова: искусственная нейронная сеть, управление с обратной связью, ошибка сети.

В последнее время существенно расширился круг задач, решаемых с помощью аппарата искусственных нейронных сетей (ИНС). В их числе – синтез систем управления, идентификация процессов, протекающих в системах, обработка данных, восстановление информации, построение расписаний и даже такие необычные применения, как создание художественных изображений.

Происходит модификация как структуры и топологии нейронных сетей, так и методов их обучения и применения. Тем не менее, нейронные сети со своим характерным поведением до сих пор остаются системами, в отношении которых существуют лишь наборы практических рекомендаций, основанные на эвристических подходах, численных экспериментах и т.п. Многие авторы, в том числе и С.Хайкин [8], указывают на то, что вопросы строгих исследований поведения нейронных сетей и построения нейросетевых решений представляют собой сложную комплексную проблему.

Любая нейронная сеть является сложной системой, которая обладает рядом специфических свойств, которые позволяют проанализировать их с позиций теории динамических систем.

Существует не так много работ, в которых нейронные сети представлены в терминах теории оптимального управления или как динамические системы.

Среди первых важно отметить ряд работ, которые посвящены построению оптимальной структуры сети. Так, в работе [5] для оптимизации структуры искусственной нейронной сети используется эволюционный алгоритм

одновременной настройки структуры весов и межнейронных связей. Одной из предпосылок, вызвавшей появление такого алгоритма, выступает сложность объективной оценки топологии сети [5, с.35-36]. Суть его заключается в следующем: кодируется как признак в генотипе ИНС число нейронов, а также величины связей, после чего вводятся операторы скрещивания, мутации и отбора. Популяции ИНС скрещиваются и отбираются с использованием усечения и элитарного отбора. Между тем, делается оговорка, что «не существует для некоторого алгоритма универсального оптимального набора параметров, позволяющего максимально эффективно решать любые поставленные задачи.» [5, с.103].

В этой же работе справедливо отмечается, что часто использование традиционных приемов обучения сетей не позволяет правильно определить компоненты выходной величины на примерах из обучающего множества, а также при необходимости работать с последовательностью выходов сети. Именно такой случай наблюдается при работе сети с адаптивной проводимостью сигнала для построения расписаний. Такая сеть не может использоваться (для ряда сигналов) информацию о значениях ошибки ни в одном слое, кроме последнего. Это также повышает актуальность поиска иного научного инструментария для управления работой такой сети. Кроме того, вне зависимости от полученной конфигурации сети, задача ее эффективного функционирования сводится к задаче ее обучения, которое должно происходить наиболее эффективным образом.

К схожей задаче можно отнести попытку построения расписания с помощью уравнивания Беллмана [10]. При этом применяется последовательное конструирование расписания с использованием принципа оптимальности любой части расписания, когда оно оптимально в целом. При формулировке задачи исходят из следующего:

Пусть T_i – длительность обслуживания i -й заявки,

W_i – время ожидания в очереди до начала обслуживания,

Игнатенков Александр Владимирович, заместитель Генерального директора по информационным технологиям ООО «Научно-технологический центр по эксплуатации железных дорог», аспирант кафедры «Автоматика, телемеханика и связь на железнодорожном транспорте». E-mail: a.ignatenkov@gmail.com

Ольшанский Алексей Михайлович, кандидат технических наук, докторант кафедры «Автоматика, телемеханика и связь на железнодорожном транспорте». E-mail: lexolshans@gmail.com

a_i - штраф за единицу времени ожидания результатов после поступления i -ой заявки,

Тогда длительность ожидания до получения результата:

$$V_i = T_i + W_i.$$

Штраф за ожидание можно записать следующим образом:

$$C = \sum_i a_i * (W_i + T_i).$$

Обозначим через D_i директивный срок решения каждой заявки. Если нам необходимо учитывать своевременное выполнение работы, то штраф будем определять по наибольшему или среднему нарушению директив сдачи i -й задачи. Критерий качества при этом определим как

$$C = \max_i a_i \{ \max[0, (V_i - D_i)] \},$$

где $\max[0, (V_i - D_i)]$ - величина превышения директивного срока.

Несоблюдение директивного срока может наблюдаться не только для конкретных задач, но и для выполнения значительной части заявок. В этом случае используется критерий, учитывающий общий штраф за нарушение директивных сроков:

$$C = \sum_{i=1}^n a_i \{ \max[0, (V_i - D_i)] \}.$$

Тогда для суммы штрафов при оптимальной очередности обслуживания первых i заявок с вышеприведенной длительностью обслуживания i -й заявки T_i уравнение Беллмана будет иметь вид:

$$C_n(T) = \min_i [C_{n-1}(T - T_i) + a_i T]. \quad (1)$$

Это выражение имеет тот смысл, что при оптимальной последовательности решения всех n задач за общее время T , последовательность решения $n-1$ задач за время $T - T_i$ также должна быть оптимальной. Тогда общий минимум штрафов достигается в той последовательности, когда перестановка обслуживания любых двух заявок не влечёт за собой уменьшение суммы штрафов.

Из (1) можно получить

$$\frac{T_{i-1}}{a_{i-1}} \leq \frac{T_i}{a_i}. \quad (2)$$

Таким образом, оптимальное расписание следует составлять, руководствуясь правилом (2), т.е. в порядке неубывания отношения длительности штрафа к штрафу за единицу времени ожидания. В случае равенства штрафов за ожидание расписание составляется по неубыванию значений T_i . Также можно доказать, что при соблюдении такого правила будет минимальным среднее число заявок в системе и средняя длительность ожидания до начала реализации заявок [11].

Недостатком данного подхода является экспоненциальный рост числа возможных перестановок при увеличении числа заявок. Кроме

того, рассмотренная задача слишком идеализирована и малоприменима для реальных задач, в особенности на железнодорожном транспорте, при составлении расписаний работы которого приходится учитывать как явные, так и неформальные ограничения. Это требует создания такого инструментария, который позволяет построить график полностью, которым и являются рассматриваемые ИНС.

Среди второй группы работ можно выделить наиболее значительные, в которых произведена попытка синтеза нейросетевого подхода и теории автоматического управления. В работе [3] идет речь о построении оптимальных временных последовательностей весов для динамической нейронной сети. В этой работе решается двухточечная краевая нелинейная задача управления, ведущая к правилам обучения сети. Оптимальная матрица функционирования сети выглядит как сумма (по всем индексам во всех выборках) внешних (векторных) произведений между вектором желаемых выходов размерности n и соответствующим выходом из определенной подсистемы, которые выделяются из системы N . Матрица весов на каждом шаге работы сети в итоге задается как оптимальная последовательность весов во времени. Авторы отмечают, что в идеальном случае весовая матрица в конечный момент времени соотносится с симметричной матрицей, эмпирически предложенной Дж. Хопфилдом для ассоциативной памяти [1].

В векторном виде уравнение эволюции нейрона записывается как:

$$\frac{dx}{dt} = -X(t) + W(t)g(x(t)), \quad (3)$$

где $W(t)$ - блочно-диагональная матрица весов, каждый блок которой задается как матрица $W1(t)$, $g(x)$ - выход нейрона.

Начальные условия задаются как составленный из нескольких выборок вектор входа.

Минимизируемый функционал качества направлен на минимизацию величины, противоположной корреляции между выходом нейрона и желаемым выходным вектором в конечный момент времени управления. В процессе движения от момента начала управления до момента окончания управления функционал штрафует некоррелированность между желаемым выходом и функцией активации нейрона.

Оптимальное управление в данном случае отыскивается как задача Лагранжа для синтеза программного управления.

Авторы отмечают, что полученные результаты относятся к частному случаю нейронной сети, которая рассмотрена в [3].

Кроме того, существуют определенные сомнения в правильности оценки скорости изменения вектора фазовых переменных состояния нейронной сети как разности между желаемым выходом и расчетным выходом сети.

Предлагаемый подход также не позволяет работать с такими нейронными сетями, которые являются не полностью управляемыми, а управляемыми только по выходу. В то время как часть сетей характеризуется только параметрами, которые отражают выход, а не внутренние переменные сети.

Работа [4] посвящена разработке интеллектуального оптимального управления с динамическими нейронными сетями. Динамическая нейронная сеть описывается стандартными уравнениями эволюции нейрона, но на структурной схеме, описывающей состояния сети, присутствуют элементы, имеющие динамические звенья, такие, как: интегрирующее звено, усилительное звено, обратные связи (отрицательные). Такие системы при определенных параметрах матриц, описывающих поведение сети, могут приводить к возникновению предельных циклов, хаотических и существенно нелинейных движений. В качестве критериев качества обучения используются квадратичные функционалы, а также абсолютные отклонения. Авторы [4] также модифицируют алгоритм оптимизации BFGS.

Направление, рассматривающее нейронные сети как самоорганизующиеся системы, представлено в работе [4]. Согласно этому подходу, рост числа нейронов в сети задается как функция от активности сети, описываемой уравнением, однородным по своему классу с (3), а также распространением сигнала во внешней среде согласно уравнению теплопроводности. Условием роста структуры нейронов является достижение аксоном заданного нейрона некоторой точки, находящейся в зоне иного нейрона, причем величина зоны определяется пользователем. «Аксон при этом растет в направлении наибольшей концентрации вещества, участвующего в его создании» [6], а внешний сигнал, поступающий извне в сеть, определяет вектор изменения активностей нейронов.

Однако данная работа не сосредоточена на процессе функционирования нейронной сети.

Достаточно близко к теме исследования подошли авторы работы [7], которые ставили задачу управления нейронной сетью как системой с запаздыванием. При этом автором налагались ограничения на веса; математическая же модель нейронной сети работы [7] родственна модели, изложенной в [3]. Как и работа [3], такая модель не полностью соответствует рассматриваемой новой топологии нейронной сети с переменной проводимостью сигнала.

Таким образом, анализ работ, посвященных синтезу управления нейронных сетей, показывает, что данная сфера является изученной и проработанной в небольшой степени.

В частности, интерес представляет задача оценки состояния многослойной искусственной нейронной сети с переменной проводимостью

сигнала [2], которая может быть использована для решения задач составления расписаний различных процессов, по наблюдаемому выходу сети или по значениям сигнала ошибки.

В этом случае мы сталкиваемся с проблемой способа формального задания эволюционных уравнений, причем возможны варианты задания их в матричном виде или в виде системы с одним входом и одним выходом.

В настоящей работе будет сделана попытка синтезировать оптимальное управление искусственной нейронной сетью с переменной проводимостью сигнала сначала как разомкнутой системой, а затем – как замкнутой.

При этом отметим, что в работе будет поставлена и решена задача Больца с квадратичным функционалом качества управления. Неприменимость линейных функционалов качества управления, а также задач на максимальное быстроедействие к управлению нейронными сетями показана в [9] и в настоящей статье рассматриваться не будет.

Рассмотрим многослойную нейронную сеть с переменной проводимостью сигнала [2], состояние которой в каждый момент времени характеризуется дифференциальным уравнением: $\frac{dE}{dt} = \sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t)) + u(E, t)$, (4)

где m – число гармоник, из которых состоит сигнал ошибки, в расчётах $m = 7$

a_i, b_i – косинус-коэффициенты и синус-коэффициенты,

w_i – частоты гармоник, определяемые спектральным анализом,

$u(E, t)$ – некоторое искомое управление.

То есть, сигнал ошибки сети во времени может быть аппроксимирован в виде совокупности гармоник.

Предположим, что существует некоторое управление u , которое мы применяем к нашей нейронной сети в уравнении (4), чтобы перевести сеть из некоторого состояния в момент времени t_0 в целевое состояние в момент t_1 .

Момент начала управления t_0 определяется моментом пересечения траекторией $E(t)$ некоторой поверхности с заданным значением уровня ошибки Δ .

Момент окончания управления t_1 является открытым и определяется в ходе решения задачи синтеза оптимального управления сетью. В содержательном смысле условие окончания управления записывается как $E(t) < \Delta$.

Синтез программного управления

Для сети, развивающейся в соответствии с (4) определим функционал качества как

$$I = \int_{t_0}^{t_1} u^2(t) dt + E(t_1) \rightarrow \min. \quad (5)$$

Для задачи (5)-(6) запишем выражение гамильтониана:

$$H(E, t, u, \Psi) = \Psi_1(t) * \left(\sum_{i=1}^m (a_i \cos(w_i t) + \right.$$

$$+ b_i \sin(w_i t)) + u(t) - u^2(t) - E(t_1), \quad (6)$$

где $\Psi_1(t)$ – вспомогательная функция.

Примем, что ограничений на управление не задано.

В соответствии с принципом максимума Л.С. Понтрягина найдем структуру оптимального управления:

$$\frac{\partial}{\partial u} H(E, t, u, \Psi) = \Psi_1(t) - 2u(t). \quad (7)$$

Покажем, что действительно достигается максимум гамильтониана по управлению:

$$\frac{\partial^2 H(E, t, u, \Psi)}{\partial u^2} = -2 < 0. \quad (8)$$

Из равенства (8) нулю найдем структуру оптимального управления:

$$u_{opt}(t) = \frac{\Psi_1(t)}{2}. \quad (9)$$

Запишем систему канонических уравнений принципа максимума Л.С. Понтрягина с учетом гамильтониана (6):

$$\begin{cases} \frac{\partial E}{\partial t} = \frac{\partial H}{\partial \Psi} = u(t) + (\sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t))) = \frac{\Psi_1(t)}{2} + (\sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t))), \\ E(t_0) = E_0, \\ \frac{\partial \Psi}{\partial t} = -\frac{\partial H}{\partial E} = -1 \end{cases} \quad (10)$$

Второе уравнение в (10) – начальные условия для уравнения (4).

Произведем проверку условий трансверсальности, чтобы получить недостающее для решения (10) условие. В общем виде условия трансверсальности с учетом (6) запишем как:

$$\delta E(t_1) - H(t_1) * \delta t_1 + \Psi_1(t_1) * \delta E = 0 \quad (11)$$

или

$$\delta E(t_1) - \left[\Psi_1(t) * \left(\sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t)) \right) + u(t) - u^2(t) - E(t_1) \right] * \delta t_1 + \Psi_1(t_1) * \delta E = 0 \quad (12)$$

Так как ограничений на значение ошибки сети в момент окончания управления не наложено, то вариации δE произвольные; выберем из них нулевые. Следовательно,

$$-\left[\Psi_1(t) * \left(\sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t)) \right) + u(t) - u^2(t) - E(t_1) \right] * \delta t_1 = 0. \quad (13)$$

Учитывая, что правый конец траектории по времени не зафиксирован, можно считать, что $\delta t_1 \neq 0$. Произведем сокращение на эту величину.

Заменив в (13) управление на его выражение (9), получим:

$$-\Psi_1(t) * \sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t)) + \frac{-\Psi_1(t)^2}{2} + \frac{\Psi_1(t)^2}{4} = E(t_1). \quad (14)$$

Решая третье уравнение из (10), получим:

$$\Psi_1(t) = -t + C. \quad (15)$$

где для различных моментов времени окончания управления с помощью (15) вычисляется значение константы C .

Тогда $u_{opt}(t) = \frac{-t+C}{2}$, подставляя которое в

первое уравнение системы (10), имеем:

$$\frac{\partial E}{\partial t} = \frac{-t+C}{2} + \left(\sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t)) \right). \quad (16)$$

Интегрируя (16), получим:

$$E(t) = \frac{Ct}{2} - \frac{t^2}{4} + \dots + \frac{a_i}{w_i} \sin(w_i t) - \frac{b_i}{w_i} \cos(w_i t) + C_1. \quad (17)$$

где C_1 вычисляется из второго условия системы (10) при известной C .

Характер уравнения (17) говорит о том, что при некотором значении t $E(t)$ будет равняться нулю, а далее стремиться к $-\infty$, что противоречит физическому смыслу ошибки сети E . Поэтому управление (15) должно остановиться при E меньше, чем наперед заданное $E(t_1)$. В силу этих заключений будем пытаться решить поставленную задачу при замкнутом управлении.

Синтез управления с обратной связью

Рассмотрим квадратичный функционал качества управления сетью, который запишем как (5), минимизирующий любое управление, прилагаемое к нейронной сети:

$$I = \int_{t_0} u^2(E, W, t) dt + E(W, t_1, u) \rightarrow \min. \quad (18)$$

Решением любой из задач управления сетью выступает оптимальное управление $u^*(E, W, t)$ с обратной связью, момент окончания управления t_1^* , а также оптимальная траектория $E^*(t, W, u^*)$ снижения ошибки под воздействием управления.

Кривые $u^*(E, W, t)$ и $E^*(t, W, u^*)$ приводят нас к способу управления, который должен быть реализован алгоритмически через элементы структуры и поведения рассматриваемой нейронной сети с переменным распространением сигнала.

Вид граничного условия (19) определяется исходя из того, что в различные моменты времени окончания управления будет существовать различный фиксированный уровень остаточной ошибки:

$$\Phi(t_1, E) = E, \quad (19)$$

где $\Phi(t_1, E)$ – функция граничного условия.

Запишем уравнение Беллмана Φ^B для задачи (4) с функционалом качества (5) и граничным условием (19):

$$\Phi^B = \frac{\partial \Phi}{\partial t} + \frac{\partial \Phi}{\partial E} * \left(\sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t)) \right) + u(E, t) - u^2(E, t), \quad (20)$$

Найдем производную функции Беллмана по управлению и приравняем к нулю для поиска структуры оптимального управления:

$$\frac{\partial \Phi^B}{\partial u} = \frac{\partial \Phi}{\partial E} - 2u(E, t) = 0. \quad (21)$$

Отсюда структура оптимального управления выражается как:

$$u(E, t) = \frac{1}{2} \frac{\partial \Phi}{\partial E}. \quad (22)$$

Подставив (22) в (20), получим

$$\frac{\partial \Phi}{\partial t} + \frac{\partial \Phi}{\partial E} * \left(\sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t)) \right) + \frac{1}{4} \left[\frac{\partial \Phi}{\partial E} \right]^2 = 0. \quad (23)$$

С учетом соображений (19) будем искать решение уравнения (23) в виде

$$\Phi(t, E) = K(t)E. \quad (24)$$

Запишем, что

$$\frac{\partial \Phi}{\partial t} = \frac{dK(t)}{dt} E, \quad (25)$$

$$\frac{\partial \Phi}{\partial E} = K(t). \quad (26)$$

Подставив (25) и (26) в (23), получим общий вид уравнения, решив которое, можно получить функцию остаточной ошибки от времени, из которой с учетом (24), (22) можно будет получить функцию управления с обратной связью:

$$\frac{E(t_2)}{E(t)} \left[\frac{\partial E}{\partial t} \right] + \frac{E(t_2)}{E(t)} \left(\sum_{i=1}^m (a_i \cos(w_i t) + b_i \sin(w_i t)) \right) + \frac{1}{4} \frac{E(t_2)}{E(t)} = 0. \quad (27)$$

При начальных условиях, т.е. $E(t_0) = E_0 = 0,8777255$ от максимальной амплитуды ряда ошибки и при выбранных основных гармониках с круговыми частотами 0,07; 1,05; 1,48; 1,7; 2,25; 2,60; 3,25; (по дан-

ным программного продукта «SCAN» и мощности амплитуды: 0,223607; 0,3; 0,3; 0,4472; 0,4472; 0,547; 0,387 от максимальной амплитуды ряда ошибки, применение метода Рунге-Кутты 4-го порядка для уравнения (27) дало следующий вид функции управления примененной нейронной сетью.

По оси абсцисс данного графика показаны отсчеты времени.

Сделанные расчёты позволяют установить следующее:

1. С ростом времени необходимая интенсивность управления сетью должна меняться также квазипериодически, согласно рис. 1.

2. Увеличение учёта числа гармоник, включаемых в уравнение эволюции ошибки сети (4), приведёт к росту точности решения.

3. В общем случае искусственная нейронная сеть не является полностью управляемой системой, однако приведенное решение позволяет считать нейронные сети с переменной проводимостью сигнала системами, управляемыми по выходу.

4. Несмотря на найденный характер управления, возникает новая проблема трансформации найденного решения в конкретные алгоритмы, изменяющие веса связей между нейронами, скорости изменения данных весов, скорость обучения сети теми средствами, которые предусмотрены в конструкции нейронной сети [2]. Для реализации полученной кривой управления необходимо провести серию численных экспери-



Рис. 1. Полученное решение в виде функции $K(t)$ (согласно (22))

ментов, устанавливающих влияние поведения различных групп весов сети на спектральный состав наблюдаемых кривых ошибки ИНС с переменной проводимостью сигнала, что является направлением развития выполняемого исследования.

СПИСОК ЛИТЕРАТУРЫ

1. Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities // Proc. Natl.Acad.Sci.USA, vol.79 pp.2554-2558, Biophysics, April 1982. 5
2. Игнатенков А.В., Ольшанский А.М. Применение искусственной нейронной сети для построения расписаний процессов на примере графика движения поездов // Современные информационные технологии и ИТ-образование. 2015. Т.2. №11. С.50-55.
3. Fahotimi, O., Dembo, A., Kailath, T. Neural network weight matrix synthesis using optimal control techniques // USA, Stanford, 1989. // Advances in Neural Information Processing Systems-2 (NIPS-2) // USA, Denver, Colorado. 8
4. Becerikli Y., Konar A.-F., Samad T. Intelligent optimal control with dynamic neural networks. // www.elsevier.com/locate/neunet; Neural Networks 16(2003), pp.251-259. 9
5. Цой Ю.Р. Нейроэволюционный алгоритм и программные средства для обработки изображений: дис. ... канд. техн. наук. Томск: Томский политехнический университет, 2007. 209 с.
6. Галимянов Ф.А., Гафаров Ф.М., Хуснутдинов Н.Р. Модель роста нейронной сети // Математическое моделирование. 2011. Т. 23. № 3. С. 101-108.
7. Андреева Е.А., Пустарнакова Ю.А. Математическая модель искусственной нейронной сети с запаздыванием // Программные продукты и системы. 2001. № 3. С. 6-9.
8. Хайкин Саймон. Нейронные сети: полный курс: Пер. с англ. 2-е издание. М.: Издательский дом «Вильямс», 2006. 1104 с.
9. Ольшанский А.М. Некоторые особенности постановки задач управления нейронными сетями // В кн.: Современные методы, принципы и системы автоматизации управления на транспорте: сборник материалов Международной научно-практической конференции (19-20 апреля 2016 г., г. Нижний Новгород). М.: Мос. гос. ун-т пут.сообщ., 2016. С.111-114.
10. Пантелеев А.В., Бортаковский А.С. Теория управления в примерах и задачах: Учеб.пособие, М.: Высш. шк., 2003. С. 487.
11. Танаев В.С., Шкурба В.В. Введение в теорию расписаний. М.: Наука, 1975.

A BOUT NEURAL NETWORK ERROR CONTROL AS AN OPTIMAL CONTROL PROBLEM

© 2016 A.V. Ignatenkov¹, A.M. Olshansky²

¹LLC "Scientific and Technological Center of Railway Implementation and Maintenance", Moscow

²Samara State Railway University

In the article the statement and solutions of optimal control problem for artificial neural with variable signal conduction are considered.

Keywords: artificial neural network, feedback control, error.

Alexander Ignatenkov, Deputy Head on Information Technologies of LLC "Scientific and Technological Center of Railway Implementation and Maintenance", Postgraduate Student of Rail Automation, Telemechanics and Communication Department of Samara State Railway University. E-mail: a.ignatenkov@gmail.com
Alexey Olshansky, Candidate of Technics, Doctoral Student of Rail Automation, Telemechanics and Communication Department of Samara State Railway University. E-mail: lexolshans@gmail.com