

АВТОМАТИЗАЦИЯ ДИАГНОСТИКИ БОЛЕЗНИ ПАРКИНСОНА: ПОСТАНОВКА ПРОБЛЕМЫ

О.В. Жвалевский (Санкт-Петербург)

Развитие информационных технологий обеспечивает медицинскую диагностику новыми инструментами и новыми возможностями для специалистов:

- современные приборы позволяют обнаруживать мельчайшие концентрации веществ, а это значит, что существовавшие до внедрения информационных технологий «грубые» шкалы могут быть существенным образом уточнены;
- современные аппаратно-программные комплексы (АПК) позволяют «погрузить» любого человека в сенсорную сеть, и тогда любая функциональная проба может быть записана (в виде полиграфической записи) и проанализирована *post factum* самым тщательным образом;
- современные вычислительные средства позволяют реализовать математическую обработку экспериментальных данных практически любой сложности и выполнить её с максимально возможной скоростью.

Эффективное применение информационных технологий, однако, невозможно без хорошо проработанной и хорошо обоснованной методики сбора и математической обработки экспериментальных данных. Если, например, имеется методика обработки кардиосигналов по Р.М. Баевскому, то эта методика, будучи реализованной в виде подпрограммы в составе какого-либо аппаратно-программного комплекса, позволяет врачу немедленно получить оценку функционального состояния пациента и дать пациенту обоснованные рекомендации. А что делать, если такой устоявшейся методики нет? Диагностика болезни Паркинсона (БП) является тем самым примером, когда такой методики нет, и основные усилия исследователей как раз направлены на создание такой методики. Если бы такая методика (для диагностики болезни Паркинсона) существовала, то её можно было бы реализовать в виде системы поддержки принятия врачебных решений и внедрить такую систему в каждодневную практику медицинских учреждений. Рассмотрим возможный подход к построению такой методики и те проблемы, с которыми нам придётся столкнуться при её построении.

Постановка проблемы

Автоматизация диагностики БП фактически опирается на две методики. Первая методика — это методика сбора и последующей математической обработки экспериментальных данных. Она широко используется на этапе исследований. Конечным результатом применения этой методики являются описания: измерительных экспериментов и требований, которым должны удовлетворять временные ряды регистрируемых физиологических показателей; эффективных методов математического анализа временных рядов, которые позволяют вычислять потенциально информативные признаки; системы распознавания, содержащие перечисление эффективных (в плане диагностики БП) критериев качества распознавания, схем применения методов машинного обучения и алгоритмов отбора информативных признаков; допустимых наборов информативных признаков; точности и достоверности всех полученных результатов. Вторая методика — это та методика, которая должна использоваться на практике и может быть реализована в виде системы поддержки врачебных решений. Неотъемлемой частью этой методики является разработанная ранее система распознавания. Таким образом, обе методики неразрывно связаны друг с другом. Более того, должна существовать возможность корректировать ранее сделанные выводы и изменять функционирующую систему распознавания как при поступлении новых данных, так и при изменении представлений о механизмах заболевания человеком БП.

С.П.Романов предложил изометрический способ регистрации тремора рук испытуемого при помощи пьезорезистивного датчика. Изометрический способ означает, что руки испытуемого находятся в неподвижном положении, а испытуемый постоянно поддерживает определённый уровень удерживаемого им усилия, при этом испытуемый следит за уровнем удерживаемого им усилия на экране компьютера, постоянно совмещая две метки, каждая из которых связана со своим датчиком (то есть со своей рукой). Тензометрический способ регистрации тремора позволяет регистрировать, по сути, только полезный сигнал, содержащий полную и точную информацию о процессах, протекающих в многокольцевой системе управления движениями по Бернштейну. Всё это существенным образом (и в хорошую сторону) отличает этот способ регистрации тремора от других способов, в особенности от тех, где используются датчики ускорения и где на вид регистрируемых треморограмм оказывают влияние антропометрические данные испытуемых и отдельные (случайные) движения отдельных звеньев локомоторного аппарата. Таким образом, у нас есть способ регистрации тремора, подходящий для диагностики БП, и исходные данные для анализа, которые мы будем называть, в соответствии с указанным способом регистрации тремора, тензотреморограммами (ТТГ). Наша основная задача заключается в построении системы распознавания, основанной на математическом анализе ТТГ. Эта система должна быть способна правильно оценивать функциональное состояние человека в шкале наименований, содержащей все возможные (допустимые) наименования от «здоров» до «болен» («болен БП»).

Построение системы распознавания существенным образом упирается в ряд проблем, которые необходимо решить при построении системы распознавания.

Первая проблема — это проблема адекватности экспериментальных данных наблюдаемым процессам и явлениям. Любые экспериментальные данные — это результат измерений. Наблюдаемые процессы непрерывны и могут быть описаны на абстрактном уровне как непрерывные функции. В противовес этому экспериментальные данные дискретны и представляются, в том числе, цифровыми сигналами, которые возникают в результате двойного квантования (квантование по времени и квантование по уровню), и шум квантования определённым, порою даже существенным образом, сказывается на качестве анализируемых сигналов. Далее, при дискретизации, мы не знаем, как проходит анализируемый процесс в промежутках между отсчётами. Если бы мы могли использовать какие-либо (модельные) предположения о том, как проходит анализируемый процесс в промежутках между отсчётами, то мы могли бы выбирать более эффективные методы математического анализа временных рядов и получать с их помощью более обоснованные результаты.

Вторая проблема — это проблема адекватности методов математической обработки самим обрабатываемым данным. Даже если данные собраны самым тщательным образом и таким же тщательным образом обработаны при помощи хорошо обоснованных методов обработки, остаётся опасность обнаружения ложных закономерностей. Дело в том, что любые выводы носят вероятностно-статистический характер и всегда основываются на каких-либо априорных предположениях (например, о виде распределений), которые скорее всего не выполняются. Применение аппарата теории вероятностей и математической статистики хотя и может быть эффективным средством анализа, на деле этот аппарат может создать у исследователя лишь видимость результата. И, наконец, самое главное — это то, что экспериментальные данные «неподвижны». Даже если речь идёт о многолетних исследованиях, когда для анализа постоянно поступают новые данные, эти данные «пассивны», воспринимаются «как есть» и обрабатываются «в совокупности». В этом смысле в экспериментальных данных, если можно так выразиться, «ничего нет», и тогда любые выводы оказываются несостоятельными, или «есть всё, что угодно», и тогда из данных потенциально могут быть извлечены любые произвольные закономерности. Что же будет, если при анализе имеющихся данных возникнут вопросы, ответы на которые в имеющихся

данных отсутствуют (то есть никак не могут быть получены из исходных данных при помощи каких-либо методов обработки)? Если бы существовала возможность провести новые измерительные эксперименты (с изменённой постановкой измерительного эксперимента) и получить новые данные, которые заведомо содержат ответы на дополнительные вопросы, то исследования можно было бы проводить наиболее эффективным образом.

Третья проблема — это проблема адекватности априорной информации. При диагностике БП основным и единственным источником априорных сведений об испытуемых являются врачи, но врачи могут поставить диагноз неправильно. Если рассматривать БП, то здесь имеются две проблемные ситуации. Во-первых, помимо собственно БП есть ещё и так называемый синдром паркинсонизма, который, вообще говоря, отличен (в плане своей этиологии) от самой болезни Паркинсона, но имеет с ней множество сходных проявлений в виде тремора, гиперкинезии и ригидности. Помимо этого, у людей может быть множество отклонений (например, психогенетического характера), которые также находят своё отражение в виде «патологического» тремора. Во-вторых, наиболее неопределённым оказывается само представление о норме. Как должна выглядеть треморограмма абсолютно здорового человека? Ситуация осложняется тем, что БП долгое время может присутствовать у человека в скрытой форме и слабым образом проявляться на треморограммах. Суть автоматизации диагностики болезни Паркинсона в том и заключается, чтобы научиться обнаруживать ранние признаки заболевания. Если бы существовал какой-нибудь решающий эксперимент (например, эксперимент, позволяющий точно определить процент допаминергических нейронов чёрной субстанции), то мы имели бы априорную информацию самого высокого качества.

Четвёртая проблема — это само понятие функционального состояния. Здоровый человек может находиться в разное время в различных функциональных состояниях, вызываемых настроением, суточной активностью, перенесённым заболеванием и т.д. и т.п. В таком же положении находится и человек, больной БП (или, в особенности, синдромом паркинсонизма). Но больной человек принимает лекарства, а это значит, что каждый приём лекарства переводит больного человека в изменённое состояние (которое, в идеале, должно быть ближе к нормальному состоянию), и тогда, чтобы получить чистую («фоновую») ТТГ для обычного патологического состояния, необходимо специальным образом отменять приём препаратов. Сам измерительный эксперимент также переводит человека в изменённое состояние. Таким образом, следует говорить о том, что есть функциональные состояния, в которых может находиться человек в разные моменты времени, а есть (функциональные) состояния, которые соответствуют состоянию человека в целом. Будем условно называть обычные функциональные состояния микросостояниями, а состояния общего вида — макросостояниями. Трудность построения системы распознавания заключается в том, что мы фактически пытаемся оценить функциональное состояние (макросостояние) по результатам обработки данных, полученных для человека, находящегося в определённом микросостоянии, причём мы не знаем заранее, что в полученных данных относится к микросостоянию, а что — к макросостоянию. Как правило, об этом разделении забывают, и это приводит к чудовищному упрощению, что обрекает исследователя на обнаружение ложных закономерностей. Если бы у исследователя была возможность учитывать скрытые микросостояния, то можно было бы получать надёжные и проверяемые результаты обработки экспериментальных данных.

Имитационное моделирование системы распознавания

Для того, чтобы построить систему распознавания, необходимо в явном виде использовать несколько моделей.

Первая модель — это модель макросостояний. Каждому макросостоянию соответствует свой набор микросостояний и система переходов между ними. Процесс перехода из одного

микросостояния в другое — это относительно быстрый переходный процесс, более всего соответствующий понятию гомеостаза, в то время как переход из одного макросостояния в другое — это относительно медленный процесс, более всего соответствующий понятию гомеокинеза. Процесс регистрации ТТГ, при котором сначала удерживается минимальный уровень усилия, а затем удерживается максимальный уровень усилия, осуществляемый также в двух различных режимах многокольцевой системы управления движениями (когда она задействована в минимальной степени и когда она задействована в максимальной степени), используется именно для того, чтобы наиболее полно проявить структуру микросостояний.

Вторая модель, которую необходимо использовать, — это модель системы, которая порождает наблюдаемые процессы. В случае БП это будет многокольцевая система управления движениями. Основная трудность построения такой модели заключается в том, что у системы есть собственные колебания, а то, что наблюдается на опыте, суть результат самого процесса регистрации, который приводит к искажению собственных колебаний. Если представить себе многокольцевую систему управления движениями, то биологическая обратная связь замыкает ещё одно кольцо, вызывая вынужденные колебания около заданного экспериментатором уровня. В различных макросостояниях характер этих вынужденных колебаний будет различным. Если бы у нас была имитационная модель многокольцевой системы управления, то мы могли бы генерировать искусственные экспериментальные данные и пропускать эти данные через систему распознавания, одновременно меняя и параметры модели, и параметры системы распознавания, подстраивая их в соответствии с некоторым функционалом качества. В предельном случае у нас должен быть набор функционирующих параллельно моделей, сравнивая которые при помощи системы распознавания, мы могли бы выбирать наилучшую в данный момент времени модель. Именно эти локальные модели и следует соотносить с функциональными микросостояниями.

Для того, чтобы наше моделирование было осуществимо, требуется использовать ещё и третью модель — модель порождения временного ряда. Действительно, если представить себе анализируемый временной ряд (ТТГ) как последовательность событий (или, точнее, систему событий, если допускать также и важную возможность перекрытия событий), то появляется возможность связать модель макросостояний и модель исследуемой системы многокольцевого управления. Действительно, с каждым макросостоянием связана своя структура микросостояний (локальных моделей многокольцевой системы управления) и своя модель порождения временного ряда (своя система событий). В свою очередь, модель порождения временного ряда неразрывно связана с предположениями о том, как протекает анализируемый процесс между отсчётами. Центральный вопрос, который здесь возникает, это вопрос о внутренней структуре анализируемых процессов: обладает ли анализируемый процесс короткой или долговременной памятью. Если одному функциональному состоянию соответствуют процессы с «длинной» памятью, а другому функциональному состоянию соответствуют процессы с «короткой» памятью, то у нас появляется хороший инструмент для автоматизированной диагностики.

В итоге мы приходим к необходимости строить не какую-то одну отдельно взятую систему распознавания (основанную на простой обработке временных рядов), а целых три взаимосвязанных и взаимоувязанных системы распознавания. На нижнем уровне располагается система распознавания событий во временных рядах. Эта система должна использовать эффективный критерий, позволяющий выделять «полезные» данные, описывающие конкретные события, и специальный алгоритм, позволяющий формировать модели порождения анализируемых временных рядов и связывать эти модели с различными микросостояниями. На промежуточном (среднем) уровне располагается система распознавания самих микросостояний, и здесь важнейшую роль играет модель многокольцевой системы управления движениями. На верхнем уровне располагается система

распознавания макросостояний. Используемая на этом уровне модель фактически описывает в обобщённом виде все виды допустимых классификаций всех участвующих в распознавании объектов на всех уровнях многоуровневой системы распознавания. Таким образом, системы распознавания, расположенные на различных уровнях, оказываются взаимосвязанными (решения задачи распознавания на одном уровне связано с решением задач распознавания на других уровнях), и это значит, что задача построения многоуровневой системы распознавания оказывается задачей многокритериальной и многоэтапной оптимизации, одновременно осуществляемой на нескольких уровнях.

Заключение

Автоматизация диагностики БП, основанная на математическом анализе ТТГ, заключается в построении многоуровневой системы распознавания, на каждом уровне которой с необходимостью используются свои модели и свои алгоритмы оптимизации. Модель порождения временного ряда, основанная на событиях, необходима для того, чтобы оптимальным образом выбирать подходящие методы обработки временных рядов, позволяющие вычислять потенциально информативные признаки. Модель многокольцевого управления, в свою очередь, необходима для того, чтобы оптимальным образом идентифицировать микросостояния. Модель макросостояний необходима для того, чтобы оптимальным образом настраивать модели нижележащих уровней, задавая для этих моделей объединяющий их внешний критерий качества. Именно эта модель сможет указать (в будущем), как именно нужно подключать имеющиеся в организме компенсаторные и восстановительные процессы, чтобы перевести человека из состояния «в патологии» в состояние «в норме». Таким образом, предлагаемый в работе системный подход к проблеме автоматизации диагностики болезни Паркинсона обладает практической значимостью для медицины.