

## Рецензия

на выпускную квалификационную работу Ивановой Елизаветы Владимировны по теме “Задачи анализа спектров тандемной масс-спектрометрии”.

Развитие методов анализа и обработки данных тандемной масс-спектрометрии является важной задачей современной протеомики. Актуальность этой задачи обуславливается высокой точностью и большим объемом данных, производимыми современными инструментами.

В работе рассмотрены две достаточно слабо связанные между собой задачи, тем не менее относящиеся к области вычислительной протеомики.

В первой части работы рассмотрена задача фильтрации и выделения значимых пиков в масс-спектрах. Проанализированы некоторые широко известные открытые алгоритмы (большая часть производителей оборудования предоставляют аналогичные алгоритмы, интегрированные в ПО прибора, которые дополнительно используют специфику конкретного прибора), показано, что практически любой “разумный” алгоритм работает лучше, чем алгоритм, предложенный как часть пакета Дерепликатор. Помимо этого предложена несложная модификация алгоритма MassSpecWavelet с целью улучшения его универсальности для работы с данными, полученными посредством современных приборов.

Во второй части работы предпринята попытка продвинуться в важной задаче кластеризации масс-спектров. Действительно, успешное решение данной задачи позволит серьезно уменьшить размеры баз данных пептидов и ускорить обработку масс-спектрометрических данных. Задача усложняется тем, что масс-спектры имеют внутреннюю “структуру”, поэтому любая используемая метрика расстояния должна учитывать эту структуру. Здесь рассмотрены некоторые известные метрики из области анализа и распознавания изображений и кажется, что они могут быть использованы при некоторых условиях для кластеризации масс-спектров.

После чтения работы остается впечатление, что автору не хватило времени на проведение полноценного исследования:

- В качестве меры качества результатов фильтрации используется очень непонятная мера “количество подозрительных пиков” вместо того, чтобы использовать аннотированные масс-спектры, для которых известен результат.
- Не очень понятно, каким образом можно использовать результаты второй главы. Позволяет ли полученные метрики близости ускорить процесс идентификации пептидов?

В свете вышесказанного считаю, что работа заслуживает оценки “хорошо”.

Рецензент Тарасов А.Л.