

О ВЫБОРЕ ФОРМАЛЬНЫХ ИЗМЕРЯЕМЫХ ПРИЗНАКОВ ВСПЛЕСКОВ Q-ТИПА ПРИ РЕШЕНИИ ЗАДАЧИ РАСПОЗНАВАНИЯ ВОЛНОВЫХ ФОРМ

В.В. Пчёлкин¹, А.Н. Пекишева²

1. Полярный геофизический институт, Апатиты, Россия

2. Кольский филиал государственного Петрозаводского университета

Аннотация. На основе существующих в литературе определений всплесков Q-типа обсуждаются возможные количественные признаки, характеризующие события данного класса. В предлагаемых пространствах признаков выполнен кластерный анализ. Обсуждается информативность предлагаемых критериев. Сформулированы два варианта решающего правила для алгоритма распознавания. Разработанные алгоритмы компьютерного поиска протестированы на экспериментальном материале.

Введение

Пример всплеска Q-типа (и фрагмента квазипериодических колебаний) показан на рис. 1. Как видно из рисунка, данное событие имеет вид затухающего колебания, т.е. характеризуется квазипериодичностью и квазиэкспоненциальным уменьшением амплитуды с течением времени. Данные волноформы впервые были описаны в работе [1] и, в дальнейшем, выделялись многими исследователями [2-7].

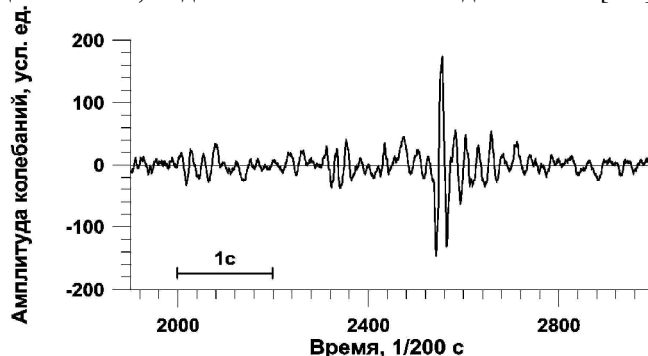


Рис. 1 Образец осциллограммы, содержащий всплеск Q-типа.

Важность автоматизированного отбора подобных, достаточно редких событий, связана с тем, что всплески позволяют оценить добротность глобального резонатора альтернативным по отношению к методам Фурье-анализа способом — по спаду амплитуды сигнала [8].

Решение задачи по распознаванию образов обычно происходит в несколько этапов (например, [9, 10]):

- 1) Формализация предметной области (определение набора признаков, характеризующих объекты).
- 2) Формирование обучающей выборки.
- 3) Уменьшение размерности пространства признаков (для предотвращения «переобучения» и улучшения точности распознавания)
- 6) Формулировка решающего правила, реализующего отбор (сортировку) в соответствии с разделяющей поверхностью в окончательно выбранном пространстве признаков.
- 5) Составление алгоритма классификации

Отметим, что приведенная последовательность является весьма условной и не включает такие важные этапы, как проверка классификатора на тестовой подборке и решение вопросов, связанных с повышением быстродействия выработанных алгоритмов.

Экспериментальные данные

Экспериментальным материалом для анализа были оцифрованные (частота квантования 200 Гц) измерения горизонтальной компоненты магнитного поля, проводимые на магнитной обсерватории Ловозеро. Методами цифровой фильтрации выделялась полоса 6-15 Гц, при помощи анализа динамических спектров и амплитудных распределений исключались записи, сопровождавшиеся техногенными помехами и магнитосферными возмущениями.

Предлагаемые пространства признаков

Изучение литературы показало, что существующие критерии отбора всплесков Q-типа носят характер

скорее качественный, морфологический, нежели количественный и сводятся к следующим признакам:

- квазипериодичность,
- превышение амплитуды колебаний определенного порога,
- закономерное (квазиэкспоненциальное) уменьшение амплитуды.

В качестве объектов классификации рассматривались выбранные фрагменты оцифрованных осциллограмм фиксированной длины с максимальным значением модуля сигнала на фиксированном месте. Для классификации нами был использован дивизимный монотетический подход. В соответствии с определениями Q-всплесков, существующими в литературе, были предложены два варианта набора признаков.

Первый вариант списка дескрипторов:

1) Признак превышения амплитуды (признак №1 в данном списке, рассчитывался по максимальному значению модуля измеряемой величины в i -м анализируемом фрагменте):

$$b_{1,i} = \max(a_{i,j}), \quad (1)$$

где i – номер фрагмента исходного ряда, j – номер измерения в оцифровке фрагмента.

2) Показатель периодичности (признак №2, рассчитывался как отношение максимальной и минимальной длительностей среди всех j -х интервалов знакопостоянства измеряемой величины в анализируемом i -м фрагменте; значение больше или равно единицы; близость к единице говорила о высокой упорядоченности смены знака измеряемой величины):

$$b_{2,i} = \frac{\max T_{i,j}}{\min T_{i,j}} \quad (2)$$

(j – номер интервала знакопостоянства во фрагменте)

3) Показатель закономерности спада амплитуды (признак №3, рассчитывался как отношение максимального и минимального во фрагменте отношения средних значений модулей измеряемой величины в соседних интервалах знакопостоянства; близость к единице говорила о закономерности спада амплитуды и наоборот):

$$b_{3,i} = \frac{\max(\text{mean}|a_{i,j}| / \text{mean}|a_{i,j+1}|)}{\min(\text{mean}|a_{i,j}| / \text{mean}|a_{i,j+1}|)} \quad (3)$$

4) Признак частоты (признак №4, рассчитывался по средней длительности интервалов знакопостоянства в i -м фрагменте).

$$b_{4,i} = \frac{1}{2 * \text{mean}(T_{i,j})}, \quad (4)$$

где j – номер интервала знакопостоянства в данном фрагменте.

5) Признак степени затухания $b_{5,i}$ (признак №5 в данном списке, рассчитывался как показатель экспоненты, наиболее хорошо – с точки зрения минимальности квадратов невязок – описывающей спад амплитуды сигнала в i -м фрагменте). Он соответствовал коэффициенту затухания в случае, если фрагмент был Q-всплеском. Этот признак формально мог быть пересчитан в признак добротности.

Второй вариант списка дескрипторов:

1) Признак превышения амплитуды (амплитудный признак – см. описание выше).

2) Признак отклонения от «идеального» всплеска на 1-й резонансной частоте.

Данный признак характеризовал отклонение формы анализируемого фрагмента (оцифрованного с частотой калибровки) от оцифрованной формы «идеального» нормированного всплеска на частоте первого шумановского резонанса:

$$a = a_0 * \sin(\omega t) * e^{-bt}$$

Отклонение формы рассчитывалось в евклидовой метрике:

$$b_{2,i} = \sum_j \sqrt{(a_{уд.,j} - a_{i,j})^2} \quad (5)$$

3) Признак отклонения от «идеального» всплеска на 2-й резонансной частоте (рассчитывался аналогично предыдущему признаку).

Обсуждение результатов

Отметим, что мы не ставили перед собой задачу строгой морфологической классификации фрагментов оцифрованных записей магнитометров в полном объеме. Нас интересовало выделение вполне определенных групп событий. Поэтому данный раздел не содержит описания морфологии всех классов, получившихся при

классифицировании. В нем рассмотрены вопросы информативности выбранных признаков, устойчивости выделяемых классов Q-всплесков, а также проблема оконтуривания.

Информативность признаков. Подтверждение резонансного характера всплесков Q-типа

Применение одного признака (например, признака квазипериодичности или признака амплитуды) при любых заданных градационных значениях не выделяет всплески Q-типа. Например, в выборке превышений их оказывалось не более 10 процентов. В выборке квазипериодических фрагментов – не более 3-х процентов (рис.2). Применение сразу двух этих признаков оказывалось более эффективным (см. ниже).

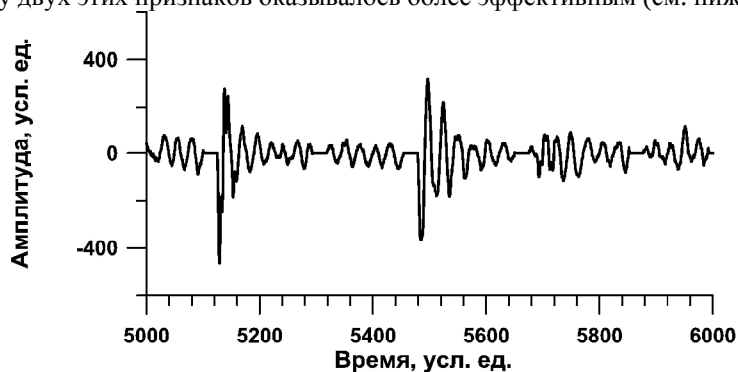


Рис. 2 Результаты отбора по признаку квазипериодичности. Присутствуют фрагменты, не являющиеся всплесками Q-типа.

На рис. 3 а, б, в показаны образцы результатов отбора по двум признакам – признаку превышения и признаку квазипериодичности на плоскости «добротность-частота» при последовательном ужесточении порогов отбора. (Добротность фрагмента рассчитывалась формально по признаку степени затухания). Как видно из рисунка 3а и 3б, все фрагменты уверенно объединяют в два кластера – соответствующие первому и второму шумановскому резонансам. Достаточно очевидно, что для данной выборки формальное применение стандартных процедур кластеризации – в частности, метода ближнего соседа или k-средних в соответствующей метрике даст 2 класса.

Усиление жесткости отбора по признакам превышения амплитуды и квазипериодичности приводит к тому, что кластер, соответствующий второму резонансу, исчезает – рис. 4в. Гистограмма рис. 5 показывает распределение событий по группам с различными градациями признака частоты (она соответствует градациям выборки рисунка 3в). Этот важный результат фактически означает подтверждение предположения о резонансном характере всплесков Q-типа. характер и связи их с первыми двумя шумановскими модами колебаний. Кроме этого, можно утверждать, что, в принципе, число признаков при отборе можно уменьшить, ограничившись всего двумя.

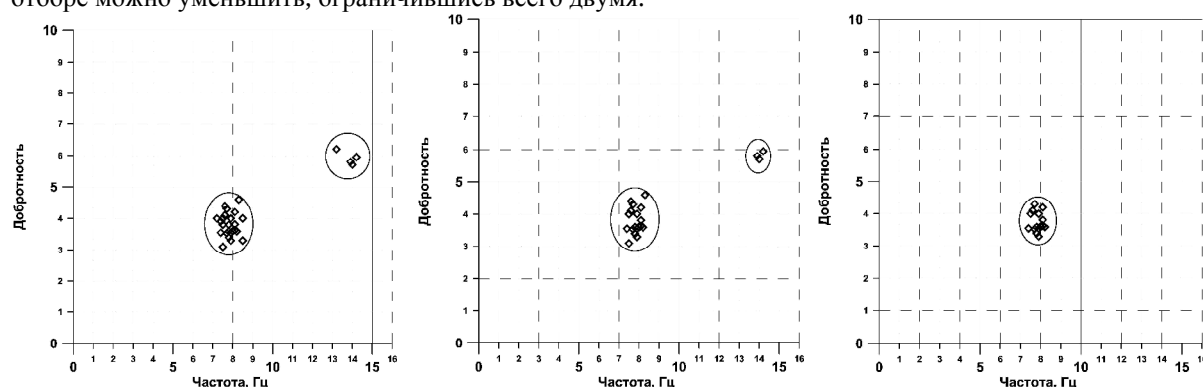


Рис. 3 Кластеры всплесков Q-типа.

Вопрос о критерии оконтуривания. Устойчивость классификации

Как известно, количественного критерия оконтуривания, как и вообще критерия правильности и объективности разбиения на классы, не существует /14-18/. Разные методы классификации и кластеризации могут выделять разные классы. А между тем эти вопросы в ситуациях, когда по отобранному материалу проводят дальнейшие количественные оценки, чрезвычайно важны. Произвол в отборе событий может отразиться на результатах этих оценок.

Косвенным доказательством объективности классификаций является совпадение состава классов при применении разных методов /14/. В нашем случае исходная выборка (длительность используемых экспериментальных рядов была от нескольких часов до суток) составляла обычно от нескольких тысяч до нескольких десятков тысяч фрагментов. Число отобранных Q-всплесков, соответственно, было от единиц до

ста (то есть порядка 0,1% исходного ряда фрагментов и менее). При этом совпадение состава классов всплесков Q-типа при различных классификациях в разных пространствах признаков было высоким, достигая 85-95%.

Вопрос о критерии оконтуривания сложнее. Однозначного его решения, повторимся, не существует. Мы решали этот вопрос на основе сопоставления увеличения дисперсии оценок добротности при ослаблении критериев отбора с величиной дисперсии оценок добротности по ярко выраженным Q-всплескам (то есть теми, которые, по мнению экспертов, являются безусловными событиями этого типа). Анализ увеличения дисперсии оценок добротности с изменением порогового значения амплитуды показал, что пороговое значение признака превышения можно несколько понизить. Это увеличивает количество отбираемых всплесков, делая их довольно регулярным материалом для оценок свойств резонатора "Земля-ионосфера". Величина максимума во фрагменте оказалась не столь критичным параметром, как предполагалось нами в начале. Более важным была «правильная» форма всплесков. (Подчеркнем, тем не менее, что в отбираемых фрагментах амплитуда превышала средние уровни фоновых шума более чем в 15 раз.)

Выводы

1) Анализ результатов классификации экспериментальных рядов показал, что всплески Q – типа образуют в плоскости «частота-добротность» устойчивый, хорошо отделенный (в смысле расстояний между объектами внутри класса и расстояниями до объектов, не вошедших в него) класс явлений.

2) Совпадение частоты объектов данного класса с частотой первой шумановской моды доказывает резонансный характер всплесков данного типа. Высокая степень совпадения состава выделяемых классов Q – всплесков говорит об объективности применяемых для выделения таких событий классификаций.

3) Показана возможность создания эффективных (то есть решающие поставленную задачу в отведенное время с заданным качеством) алгоритмов автоматизированного отбора событий данного типа в результатах измерений шумового магнитного поля КНЧ-диапазона.

4) Автоматизированный отбор всплесков Q-типа показал, что они встречаются более часто, чем считалось ранее и могут являться достаточно регулярным материалом (обычно – несколько событий в час и несколько десятков за сутки) для оценок свойств глобального резонатора "Земля-ионосфера". Было показано, что пороговое значение критерия превышения амплитуды можно понизить (при условии сохранения жестких требований к форме нормированного фрагмента).

Работа выполнена при поддержке Программой фундаментальных исследований Отделения физических наук РАН № 4.11 "Электродинамика атмосферы, радиофизические методы исследований атмосферных процессов" (проект № 4.5 "Атмосферное электричество в нижней атмосфере полярных широт").

Список литературы

- 1) T. Ogawa, Y. Tanaka, T. Miura, M. Yasuhara – Observations of Natural ELF and VLF Electromagnetic Noises by using Ball Antennas. // J. of Geomagnetism and Geoelectricity, vol.18, № 4, pp. 443 – 454, 1966.
- 2) П.В. Блюх, А.С. Николаенко, Ю. В. Филиппов – Глобальные электромагнитные резонансы в полости земля-ионосфера. // Киев, «Наукова думка», 1977.
- 3) Александров М.С. и др. - Флуктуации электромагнитного поля Земли в диапазоне СНЧ. // Коллективная монография. Изд-во «Наука», 1972, 1-100.
- 4) Lones D. Ll., Kemp D. T. – The Nature and Average Magnitude of the Sources of Transient Excitation of Schumann Resonances. // JATP, 1971, 33, p.557-566.
- 5) T. Ogawa., Tanaka Y., Fraser-Smith A.C., Gendrin R. – Worldwide Simultaneity of a Q-Type ELF Burst in the Schumann Resonance Frequency Range. // J. of Geomagnetism and Geoelectricity, vol.19, № 4, pp. 377 – 384, 1967.
- 6) T. Ogawa, K.Kozai, H.Kawamoto, M. Yasuhara, A. Huzita – Shumann resonances observed with a balloon in the stratosphere // JATP, Vol. 41, pp. 135-142, 1979.
- 7) С.М. Крылов, Н. П. Владимиров – Некоторые характеристики электромагнитных резонансов полости Земля-ионосфера. // Геомагнитные исследования, № 9, 1967, с.80-88.)
- 8) Пчелкин В.В., Белоглазов М. И., Васильев А. Н., Воронин А. И. - Всплески Q-типа в магнитных колебаниях КНЧ диапазона как явление, отражающее свойства глобального резонатора Земля-ионосфера // Геомагнетизм и аэрономия, 2010, Т. 50, N 5, с. 1-9.
- 9) Вапник В.Н., Червоненкис А.Я. Теория распознавания образов (статистические проблемы обучения). – М.:Наука, 1974.-415 с.
- 10) Загоруйко Н.Г. Методы распознавания и их применение.-М. : Сов.радио, 1972. -206 с.