УДК 543.552.054.1

ИДЕНТИФИКАЦИЯ СЕРИИ АНАЛИТИЧЕСКИХ СИГНАЛОВ НА ПРИМЕРЕ ОПРЕДЕЛЕНИЯ НЕКОТОРЫХ МЕТАЛЛОВ МЕТОДОМ ИНВЕРСИОННОЙ ВОЛЬТАМПЕРОМЕТРИИ

Кузнецов В.В., Романенко С.В.

Томский политехнический университет, Томск, e-mail: kuvv@sibmail.com

В вольтамперометрии, большую роль играет выделение полезного сигнала в серии вольтамперных кривых, снятых в одних и тех же условиях. В силу влияния на сигнал различного рода помех, в процессе анализа вольтамперные кривые могут иметь ложные пики, которые необходимо исключить из дальнейшего рассмотрения, а также идентифицировать нужные пики и привязать их к химическим элементам. От правильности привязки пиков к химическим элементам напрямую зависит результат массовой доли вещества в анализе, вычисляемой на основе метода «введено – найдено». В настоящий момент, процесс идентификации пиков ложится на плечи лаборанта, который визуально оценивает, какие пики к каким химическим элементам относятся, что бывает довольно сложно и зависит от его квалификации. В статье предложен автоматизированный алгоритм привязки аналитических сигналов в форме пиков к химическим элементам. Алгоритм является устойчивым к дрейфу центров пиков как относительно друг друга, так и относительно электрохимических потенциалов определяемых элементов. Все параметры модели имеют физических смысл и являются константами для конкретной линейки вольтамперометрических анализаторов. Алгоритм был апробирован на данных, полученных с помощью прибора ТА—07, фирмы ООО «НПП Техноаналит», г. Томск и реализован в программе Valab Professional, поставляемой в комплекте с этим прибором.

Ключевые слова: инверсионная вольтамперометрия, пик, метод добавок, кластеризация, идентификация

IDENTIFICATION OF A SERIES OF ANALYTICAL SIGNALS ON THE EXAMPLE OF THE SOME METALS DETERMINATION BY THE METHOD OF INVERSION VOLTAMMETRY

Kuznetsov V.V., Romanenko S.V.

Tomsk Polytechnic University, Tomsk, e-mail: kuvv@sibmail.com

The extraction of the desired signal in the series of voltammetric curves, plotted in the same conditions, plays an important rolein voltammetry. By virtue of effect on the signal of various noise during the analysis the voltammetric curves can have spurious peaks which must be excluded fromfurther consideration, the real peaks should be identified and they must be bound to the chemical elements. The result of mass fraction of substance in the analyte, obtained by the method of «introduced – found» dependsdirectlyon the correctness of binding peaks to the chemical elements. At present, thelaboratory assistant is responsible for process of peaksidentification who estimates visually which peaks correspond to which chemical elements that it is quite difficult to do and this estimation depends on the qualificationofthelaboratory assistant. This paper proposes the automated algorithm of the analytical signals binding in the peak shape to the chemical elements. The algorithm is robust to the drift of the peak centersas relative to each another and as relative to the electrochemical potentials of the determined elements. All parameters of the model have the physical meaning and they are the constants for the specific voltammetric analyzers. The algorithm was tested on the data, obtained by the analyzer TA-07 of theTomsk company OOO «NPP Tehnoanalit», and implemented in the program Valab Professional, supplied with this device.

Keywords: voltammetry, peak, addition method, clustering, identification

Привязка аналитических сигналов в форме пиков на вольтамперных кривых к необходимым химическим элементам является ключевой задачей при автоматической обработке вольтамперных кривых. Например, при учёте линии остаточного тока под пиком [2, 3] необходимо знать, какому именно химическому элементу он принадлежит. Также принадлежность пиков к химическим элементам актуальна в методах разделения перекрывающихся пиков [4].

Сложность поставленной задачи заключается в том, что полученные пики на вольтамперограмме могут дрейфовать случайным образом как относительно друг друга, так и относительно теоретических электрохимических потенциалов химических элементов.

Целью статьи является разработка методов и алгоритмов привязки пиков вольтамперограммы к конкретным химическим элементам, основанных на объединении пи-

ков отдельных серий вольтамперограмм в кластеры и дальнейшей привязки кластеров к соответствующим химическим элементам (идентификация кластеров).

На рис. 1 показана типичная картина вольтамперных кривых, полученная при определении массовых концентраций Zn, Cd, Pb, Cu в реальных пробах питьевой воды по методике [5].

На рисунке изображены три серии вольтамперных кривых, характеризующие (снизу вверх) фон, пробу и пробу с добавкой (в каждой серии по 5-9 кривых). На каждой кривой маркерами обозначены максимумы пиков, которые объединены в кластеры (обозначены прямоугольниками).

Вертикальными пунктирными линиями обозначены теоретические потенциалы определяемых химических элементов. Сплошные линии показывают связь кластеров с химическими элементами.

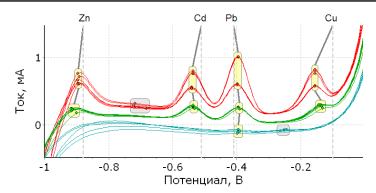


Рис. 1. Серия вольтамперограмм фона, пробы и пробы с добавкой, с привязкой пиков к химическим элементам

При анализе проб аналитик вручную выставляет соответствие пика и химического элемента, что сказывается на времени анализа. Сложность автоматизации привязки пиков к химическим элементам заключается во множестве факторов, участвующих в процессе анализа:

- пики вольтамперограмм находятся не строго друг под другом, а дрейфуют на небольшом участке. Такое поведение вызвано как химическими процессами, так и особенностями электронной части анализатора. Размах дрейфа не превышает заданную величину, которая для конкретного вольтамперометрического анализатора постоянна. Обозначим этот размах как Hd;
- вольтамперограммы могут иметь ложные пики (на рис. 1 эти пики обозначены кластерами, не связанными с химическими элементами). Ложные пики образуются за счёт помех, которые могут иметь различный частотный характер и по виду напоминать пик.

Также максимумы пиков в сериях могут быть смещены от теоретического положения химических элементов (на рис. 1 пики в сериях соединены сплошной линией с те-

оретическими расположениями элементов, обозначенных пунктирной линией). Вклад в смещение обычно даёт потенциал хлоридсеребряного электрода, дрейф которого не должен превышать 100 мВ. Обозначим максимальное смещение пиков как *Hs*.

Решение задачи автоматизации привязки пиков к химическим элементам состоит из двух этапов:

- кластеризация пиков;
- идентификация кластеров.

Под кластеризацией подразумевается объединение пиков в кластеры (группы) по степени их близости друг к другу для каждой серии кривых фона, пробы и пробы с добавкой.

Идентификация — это привязка кластеров к химическим элементам.

На рис. 2 показан схематичный пример найденных пиков у серии из 5 вольтамперограмм (обозначены точками). Жирной линией обозначена усреднённая вольтамперограмма.

Найденные пики условно сгруппированы в кластеры (обозначенные прямоугольниками) со своим номером.

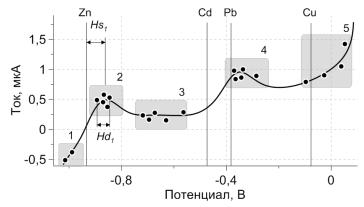


Рис. 2. Усреднённая вольтамперная кривая с кластерами пиков

Через Hs_{I} обозначен дрейф первого кластера относительно электрохимического

потенциала Zn, а через Hd_{I} обозначен размах пиков внутри первого кластера. В об-

щем случае необходимыми условиями для привязки кластера к химическому элементу являются:

$$Hs_i < Hs$$
, $Hd_i < Hd$

Если для какого-либо кластера не выполняется одно из этих условий, то кластер считается ложным.

Кластеризация пиков

Для кластеризации пиков необходимо проанализировать их потенциалы по признаку отдалённости друг от друга. На сегодняшний день существует множество алгоритмов кластерного анализа [1]. Все они являются итерационными и рассчитаны для многомерных систем.

В нашем случае мы имеем всего одну размерность и явно сгруппированные элементы пиков, что позволяет существенно упростить алгоритм. Разделение на кластеры осуществляется в два этапа:

- построение дендрограммы [1];
- поиск кластеров в дендрограмме.

Построение дендрограммы осуществляется по данным потенциалов центров пиков для каждой серии вольтамперограмм. Пространством признаков будут являться центры пиков, а количество уровней дендрограммы (число шагов слияния) определяется параметром Hd. Слияние происходит до того момента, пока разница между сливаемыми элементами будет меньше Hd.

На рис. 3 показана дендрограмма пиков графика с рис. 2 с обозначением кластеров $(c_1 \dots c_s)$.

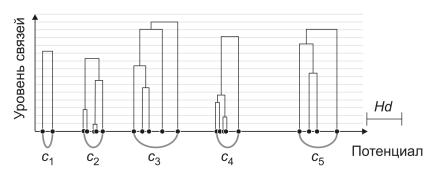


Рис. 3. Дендрограмма центров пиков

Для устранения ложных кластеров необходимо проверить условия:

$$Wc_{i} < Hd,$$

$$n - 2 \le Nc_{i} \le n,$$
(1)

где Wc_i – ширина кластера i, Nc_i – число элементов в кластере i, n – число вольтамперограмм в серии (в данном примере их 5).

Параметр Nc_i может колебаться от 1 до n. Если Nc < n, то либо имеют место «выпавшие» вольтамперограммы (искаженные, сильно зашумлённые и т. д. вследствие внешних факторов), либо пики настолько малы, что у некоторых вольтамперограмм в серии они отсутствуют (имеет место для вольтамперограмм фона). В обоих случаях число таких пиков невелико (если анализатор исправен и эксперимент выполняется в соответствии с методикой) и обычно не превышает одного-двух.

В таблице показаны значения признаков для каждого из кластеров (рис. 3). Из таблицы видно, что условиям удовлетворяют только кластеры c_2 и c_4 (тёмные области соответствуют неудовлетворительным условиям).

Значения признаков для изолированных областей гистограммы

| Признак / кластер | $\mathbf{c}_{_{1}}$ | c_2 | c ₃ | c ₄ | c ₅ |
|-------------------|---------------------|-------|----------------|----------------|-----------------------|
| Wc | ı | - | X | ı | X |
| Nc | 2 | 5 | 5 | 5 | 4 |

Для методики [5] и анализатора ТА-07 производства ООО «НПЦ Техноаналит», г. Томск, Hd = 30 мкВ.

Идентификация кластеров

Полученные на предыдущем шаге кластеры необходимо привязать к соответствующим химическим элементам (в дальнейшем, к шаблону). В общем случае:

- число кластеров может быть как больше, так и меньше числа элементов в шаблоне;
- \bullet кластеры могут быть смещены относительно элементов шаблона (см. параметр Hs).

Идентификация кластеров проводится путём формирования двух матриц (рис. 7):

- матрица соседей;
- матрица сходства.

На рис. 4 для примера формирования матриц изображён вариант элементов шаблона (кружки) и кластеров (линии).

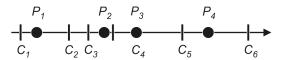


Рис. 4. Пример расположения кластеров и элементов шаблона

Матрица соседей определяет расстояние L от кластера до ближайших соседних элементов шаблона (рис. 5).

Причём из матрицы соседей должны быть исключены строки, в которых расстояние *L* превышает величину *Hs*. Таким образом, остаются только те связи, которые удовлетворяют дрейфу хлоридсеребряного электрода. Как уже было сказано, значение параметра *Hs* находится в окрестности 100 мВ. Занижение этого параметра приведёт к исключению «хороших» пиков, что может привести к увеличению погрешности расчёта концентрации, а завышение параметра приведёт к неустойчивости модели и, соответственно, неверной привязке пиков к химическим элементам.

| C | P | $\mid L$ | |
|-------|---------|----------|--|
| C_I | P_{I} | L_{I} | |
| C_2 | P_I | L_2 | |
| C_2 | P_2 | L_3 | |
| | | | |
| C_6 | P_4 | L_8 | |

Рис. 5. Демонстрация построения матрицы соседей

В матрице соседей могут содержаться варианты, когда один элемент шаблона присвоен двум кластерам (например, P_1 ссылается на C_1 и C_2) и/или когда один кластер ссылается на два шаблона (например, C_2 ссылается на P_1 и P_2).

² Для исключения таких ситуаций необходимо преобразовать матрицу соседей к матрице соответствий. Для этого необходимо пересчитать столбец L:

$$L_i = |L_i - Lm|$$

где, Lm — медиана столбца L. Далее необходимо удалить строки дублирующихся элементов шаблона и кластеров с максимальным значением L_i .

Медиана в данном случае определяет тенденцию к похожему смещению всех элементов относительно шаблона. Такой выбор критерия удаления дублирующихся элементов связан с тем, что основной вклад в смещение пиков вольтамперограмм даёт хлоридсеребряный электрод, который не меняется в процессе всего эксперимента, что приводит к примерно одинаковому смещению пиков для всех серий.

После применения алгоритма для рис. 6 останутся связи (рис. 4):

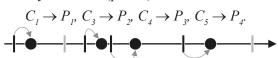


Рис. 6. Демонстрация матрицы соответствий

Кластеры C_2 и C_6 будут ложными (на рис. 10 выделены серым цветом).

Отдельно выделим случай, когда присутствует один кластер либо один элемент в шаблоне. Тогда привязка идёт по минимальному расстоянию либо от кластера, либо от шаблона соответственно без построения матриц.

Выводы

Разработанный способ на основе алгоритмов кластеризации и матрицы соответствий позволяет осуществить привязку пиков вольтамперограмм к химическим элементам, что является ключевой задачей при автоматической обработке вольтамперных кривых.

Алгоритм устойчив к дрейфу центров пиков как относительно друг друга, так и относительно электрохимических потенциалов определяемых элементов.

Входным данным алгоритма служат всего два параметра *Hd* и *Hs*, несущих физический смысл и являющихся константами для конкретной линейки анализаторов.

Предложенный способ идентификации является универсальным алгоритмом привязки массива кластеров к массиву шаблонов и может применяться в любой другой области с одномерным распределением данных.

Список литературы

- 1. Енюкова И.С. Факторный, дискриминантный и кластерный анализ. Пер с англ. Ч.У.Мьюллер, У.Р.Клекка и др. М.: Финансы и статистика, 1989. 215 с.
- 2. Романенко С.В., Романенко Э.С., Колпакова Н.А. Применение сплайн-функции дробной степени для опи-

сания базовой линии при определении платины методом инверсионной вольтамперометрии // Журн. аналит. химии. $2001.~T.~56,~N\!\!\!\! \ge 1.~C.~60-64.$

- 3. Романенко С.В., Ларин С.Л., Ларина Л.Н. Методика компенсации систематической погрешности учета базовой линии на примере определения свинца и меди методом инверсионной вольтамперометрии // Изв. вузов. Химия и хим. технология. 2002. Т. 45, № 3. С. 81 84.
- 4. Романенко С.В., Шеховцова Н.С., Карачаков Д.М. Развитие метода деления сигналов (SRRM) для разрешения перекрывающихся инверсионно-вольтамперометрических пиков // Известия Томского политехнического университета. 2008. Т. 312. № 3. Химия. С. 48 53.
- 5. МУ 31-03/04. Количественный химический анализ проб природных, питьевых и сточных вод. Методика выполнения измерений массовых концентраций цинка, кадмия, свинца и меди методом инверсионной вольтамперометрии на анализаторах типа ТА.

References

- 1. Eniukova I.S. Faktornyi, diskriminantnyi i klasternyi analiz. Per s angl. Ch.U.M'iuller, U.R.Klekka i dr. M.: Finansy i statistika, $1989.-215~\rm p.$
- 2. Romanenko S.V., Romanenko E.S., Kolpakova N.A. Primenenie splain-funktsii drobnoi stepeni dlia opisaniia bazovoi linii pri opredelenii platiny metodom inversionnoi vol'tamperometrii. Zhurn. analit. khimii. 2001. T. 56, no. 1. PP. 60 64.

- 3. Romanenko S.V., Larin S.L., Larina L.N. Metodika kompensatsii sistematicheskoi pogreshnosti ucheta bazovoi linii na primere opredeleniia svintsa i medi metodom inversionnoi vol'tamperometrii // Izv. vuzov. Khimiia i khim. tekhnologiia. 2002. T. 45, no. 3. PP. 81 84.
- 4. Romanenko S.V., Shekhovtsova N.S., Karachakov D.M. Razvitie metoda deleniia signalov (SRRM) dlia razresheniia perekryvaiushchikhsia inversionno-vol'tamperometricheskikh pikov // Izvestiia Tomskogo politekhnicheskogo universiteta. 2008. T. 312. no. 3. Khimiia. PP. 48 53.
- 5. MU 31-03/04. Kolichestvennyi khimicheskii analiz prob prirodnykh, pit'evykh i stochnykh vod. Metodika vypolneniia izmerenii massovykh kontsentratsii tsinka, kadmiia, svintsa i medi metodom inversionnoi vol'tamperometrii na analizatorakh tipa TA.

Рецензенты:

Сечин А.И., д.т.н., профессор, Федеральное агентство по образованию, НИУ РЭТ Томский политехнический университет, г. Томск;

Гольдштейн А.Е., д.т.н., профессор, заведующий кафедрой информационно-измерительной техники Томского Политехнического Университета, г. Томск.

Работа поступила в редакцию 24.06.2014.