

## **РАЗРАБОТКА АЛГОРИТМА ГЕНЕРАЛИЗАЦИИ ТОПОГЕОДЕЗИЧЕСКИХ ДАННЫХ СЕЙСМОРАЗВЕДКИ**

**Загайнова С.С.**

**Научный руководитель – Картушинский А.В.\***

**Научный руководитель – Ходяев А.В.\*\***

*Сибирский федеральный университет*

*Приводится описание алгоритма генерализации пространственных данных результатов сейсмических исследований геологических объектов. Проведен анализ и сравнение с существующими алгоритмами. Представлены результаты асимптотического эксперимента и эксперимента по использованию процессорного времени в ходе оценки эффективности разработанного алгоритма.*

*Ключевые слова: генерализация топогеодезических данных, сейсморазведка, ГИС*

### **Введение**

Топогеодезические данные сейсморазведки – это положение пунктов возбуждений и приема сигнала в пространстве. Эти данные используются для анализа изученности сейсморазведкой территории и при проектировании новых сейсмических исследований. Планово-высотные данные используются для привязки аппаратных измерений к местности. Основная обработка таких данных проводится с использованием геоинформационных систем (ГИС).

Геоинформационные системы могут передавать информацию по сети Интернет. Пользователи удаленно получают данные различного формата (изображения, текстовую информацию и т.д.). Топологические данные сейсморазведки представляют собой огромный массив информации. В условиях низкой скорости передачи данных, в сравнении с настольными ГИС, встает задача сокращения объема выводимых данных без потери их информативности. Для решения этой задачи используются алгоритмы генерализации данных или обобщения позиционных данных о пространственных объектах.

### **Существующие алгоритмы генерализации**

В литературе представлено несколько алгоритмов генерализации данных. Наиболее распространенными являются алгоритмы таких авторов как Dettori и Falcidieno, Douglas и Peucker, а также Barbara P. Buttenfield.

Эти алгоритмы используют механизм управления количеством точек, основанный на значении допустимой ошибки. Используется числовой параметр, который определяет ширину порога. Точки, попавшие в диапазон порога, удаляются.

Если параметр будет слишком велик, то в процессе обработки данных могут быть удалены критические точки, а это означает, что геометрия пространственных объектов (сейсмических съемок) будет изменена.

Hoffman и Richards предложили, что линии должны быть сегментированы в пунктах минимального искривления. Другими словами, пункты минимального искривления – критические точки.

Основным требованием к процедуре генерализации пространственных объектов сейсмических съемок является сохранение критических точек – точек излома (перегиба линии). Алгоритмы Dettori и Falcidieno, Douglas и Peucker, Barbara P. Buttenfield не выполняют данное требование.

Для доказательства рассмотрим наиболее часто используемый в ГИС алгоритм Douglas и Peucker.

По алгоритму Douglas и Peucker пользователь вводит значение допустимой ошибки  $\epsilon$ , соединятся первая и последняя точка прямой, в обе стороны от этой прямой строится коридор, ширина которого равна допустимой ошибке  $\epsilon$ . Затем сравнивается расстояние от каждой точки до прямой с введенным значением ошибки  $\epsilon$ :

- если расстояние от точки до прямой больше, чем допустимая ошибка  $\epsilon$ , то точка остается;
- если расстояние от точки до прямой меньше, чем допустимая ошибка  $\epsilon$ , то точка удаляется.

Другими словами, те точки, которые попали в коридор, удаляются, а те, которые оказались за его пределами – сохраняются.

На рис. 1 представлен пример генерализации линии по алгоритму Douglas и Peucker.

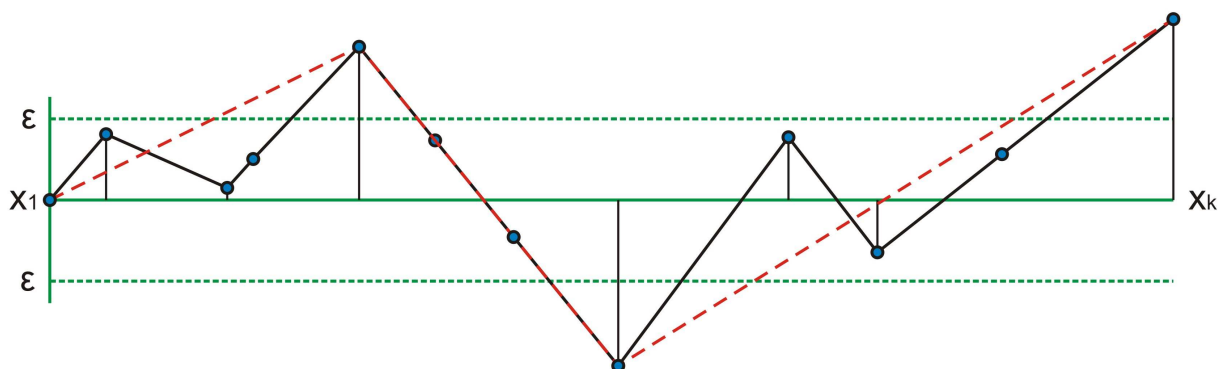


Рис. 1. Генерализация линии по алгоритму Douglas и Peucker

Алгоритм Douglas и Peucker не учитывает важность критических точек в картографии. Данные после обработки по этому алгоритму недостаточно сохраняют информативность, не все критические точки отображены для графического представления, ввиду чего, геометрия линии искажена и недостаточно точна.

Важным условием алгоритма генерализации данных – сохранение информативности.

Необходима разработка алгоритма, способного выполнять генерализацию пространственных объектов без потери критических точек.

### Разработка алгоритма

Предложенный алгоритм генерализации данных включает в себя три этапа:

1. Расчет расстояний между точками.
2. Сравнение расстояний с радиусом окружности, построенной вокруг каждой точки (значение радиуса окружности задается пользователем).
3. Обнаружение и восстановление критических точек.

Расчет расстояний между точками осуществляется по теореме Пифагора. Затем строится окружность радиуса  $R$  вокруг каждой точки. Радиус  $R$  равен желаемому расстоянию между точками для отображения. Точки, попавшие в окружность, удаляются.

Радиус окружности  $R$  – это параметр, от которого зависит, на каком расстоянии друг от друга будут отображаться точки. В сейсморазведке расстояние между точками взрыва и приема колебаний могут отличаться. На рис. 2 показано, как при помощи окружности радиуса  $R$  сокращается количество выводимых в ГИС данных.

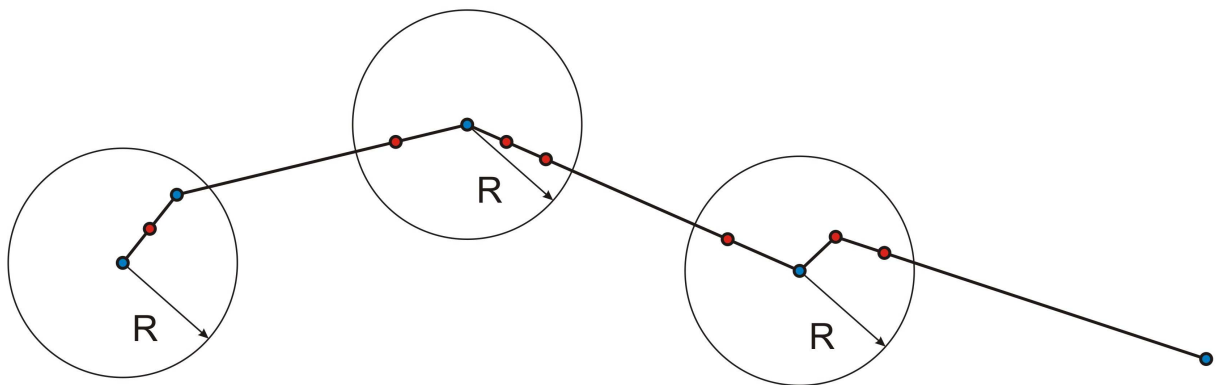


Рис. 2. Удаление точек, попавших в радиус окружности

Во время удаления точек, попавших в радиус окружности, также удаляются и критические точки – точки перегиба линии. После этого геометрия линии меняется (рис. 3, рис. 4). Такая форма сейсмического профиля не может быть использована для дальнейшей работы.

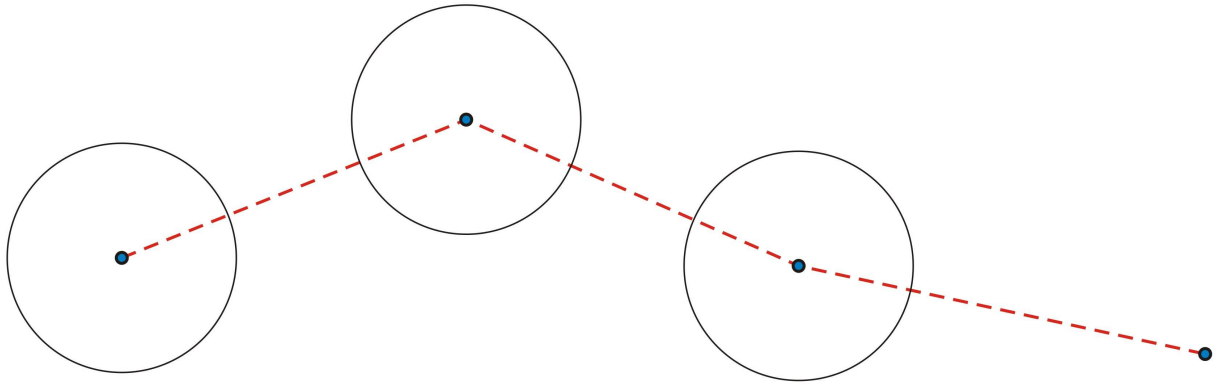


Рис. 3. Форма линии после удаления критических точек

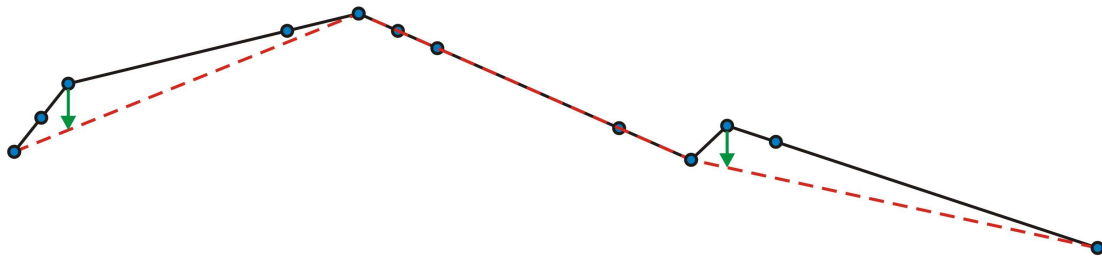


Рис. 4. Начальная и измененная геометрия линии

Далее происходит восстановление критических точек линии. Процедура восстановления использует разницу углов текущего и предыдущего сегментов в качестве критерия определения точек перегиба. Восстановленная геометрия линии содержит меньшее количество точек (рис. 5).

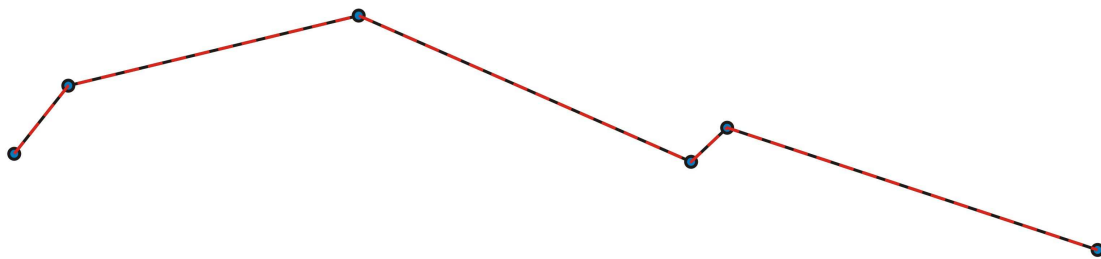


Рис. 5. Геометрия линии после генерализации

Пользователь может задавать значение угла. Если необходимо чтобы графическое представление данных было более информативно, тогда оптимальным значением угла считается  $0.01^\circ$ . Если не требуется высокой информативности, тогда оптимальным считается значение угла  $0.1^\circ$ .

### Оценка эффективности алгоритма

Для определения эффективности алгоритма проведен асимптотический эксперимент, а также эксперимент по использованию процессорного времени. Для сравнения взят алгоритм Douglas и Peucker.

Реальное время выполнения каждого отдельного шага зависит от конкретного вычислительного устройства.

В табл. 1 приведен перечень операций для предложенного алгоритма и для алгоритма Douglas и Peucker.

Таблица 1

### Перечень операций алгоритма

Предложенный алгоритм	Алгоритм Douglas и Peucker
<ol style="list-style-type: none"> <li>1. Вводится радиус окружности <math>R</math>.</li> <li>2. Расчет расстояния между точками <math>l</math>.</li> <li>3. Сравнение <math>R</math> и <math>l</math>: <ul style="list-style-type: none"> <li>– если <math>R &lt; l</math>, точка остается;</li> <li>– если <math>R &gt; l</math>, точка удаляется.</li> </ul> </li> </ol>	<ol style="list-style-type: none"> <li>1. Вводится допустимая ошибка <math>\epsilon</math>.</li> <li>2. Построение коридора.</li> <li>3. Соединение прямой линией точек <math>x_1</math> и <math>x_n</math>.</li> <li>4. Расчет расстояния <math>l</math> от точки <math>x_i</math> до прямой между точками <math>x_1</math> и <math>x_n</math>.</li> <li>5. Сравнение <math>\epsilon</math> и <math>l</math>:</li> </ol>

<p>4. Построение прямой между первой и третьей точками:</p> <ul style="list-style-type: none"> <li>– если вторая точка лежит на прямой, то она удаляется;</li> <li>– если вторая точка не лежит на прямой, тогда она является критической.</li> </ul>	<ul style="list-style-type: none"> <li>– если <math>l &lt; \varepsilon</math>, точка удаляется;</li> <li>– если <math>l &gt; \varepsilon</math>, точка остается.</li> </ul>
---	---

Количество операций алгоритма Douglas и Peucker представлено выражением (1), а количество операций предложенного алгоритма генерализации выражением (2).

$$K = 1 + 1 + 1 + i + i = 2i + 3 \quad (1)$$

$$K = 1 + (i - 1) + (i - n) + (i - 2) = 3i - 2 - n \quad (2)$$

где  $K$  – количество операций,

$i$  – количество точек,

$n$  – количество исключенных точек.

В ходе асимптотического эксперимента установлено, что на выполнение предложенного алгоритма требуется  $3i - 2 - n$  операций, а на выполнение алгоритма Douglas и Peucker требуется  $2i + 3$  операций.

Давая оценку быстродействия алгоритма, следует рассмотреть поведение вычислительного процесса в среднем и, отдельно, в экстремальных для него условиях, то есть – в худшем случае.

В худшем случае, количество операций в предложенном алгоритме равно  $3i - 2$ , в алгоритме Douglas и Peucker  $2i + 3$ .

В табл. 2 приведены данные процессорного времени, затраченного на каждый алгоритм при различных наборах данных. В качестве исходных данных использовались три ломаные линии, состоящие из 10, 50 и 100 точек. При расчете времени предложенного алгоритма использовались три значения радиуса: 3 м, 13 м и 30 м. Значения для  $R$  и  $\varepsilon$  выбирались таким образом, чтобы количество точек линий после генерализации было идентичным.

Таблица 2

## Затраченное процессорное время

	Предложенный алгоритм			Алгоритм Douglas и Peucker		
Количество точек линии: 10						
	$R$			$\mathcal{E}$		
	3	13	30	0.5	4	6
$t$ , мс	31	16	15	46	31	20
Количество точек линии: 50						
	$R$			$\mathcal{E}$		
	3	13	30	9.5	79	91
$t$ , мс	94	63	47	118	85	62
Количество точек линии: 100						
	$R$			$\mathcal{E}$		
	3	13	30	25	165	196
$t$ , мс	172	94	62	206	114	78

## Выводы

Большинство алгоритмов не учитывают важность критических точек при обработке топогеодезических данных сейсморазведки. Геометрические данные после генерализации недостаточно сохраняют требуемую информативность. Предложенный алгоритм генерализации учитывает критические точки линий сейсмических съемок, имеет меньшее количество операций по сравнению с другими алгоритмами и затрачивает меньшее количество процессорного времени.