El diseño del algoritmo TRACLUS se centra en el análisis y agrupación de trayectorias mediante un enfoque basado en la densidad. Este proceso se divide en varias etapas, que se describen a continuació:

**1. Definición de la Función de Distancia**

La primera etapa del algoritmo TRACLUS implica la definición de una función de distancia compuesta que permite medir la similitud entre trayectorias. Esta función se basa en tres componentes principales:

* **Distancia Perpendicular (d⊥):** Mide la distancia más corta entre un punto y una trayectoria, considerando la línea que pasa por el punto y es perpendicular a la trayectoria.
* **Distancia Paralela (d<sub>k</sub>):** Evalúa la diferencia en longitud a lo largo de la trayectoria entre dos puntos.
* **Distancia Angular (dθ):** Calcula la diferencia en la orientación entre segmentos de trayectorias.

Estas distancias se calculan para cada par de puntos en las trayectorias y luego se ponderan para obtener una medida de distancia compuesta, que refleja la disimilitud total entre las trayectorias.

**Implementación:** Se calculan estas distancias para cada par de segmentos de línea. La combinación ponderada de estas medidas proporciona una métrica compuesta que se utiliza para evaluar la similitud general entre segmentos, permitiendo una comparación detallada que tiene en cuenta tanto la posición como la orientación.

**2. Identificación de Puntos Significativos**

Una vez definida la función de distancia, el siguiente paso consiste en identificar los puntos más significativos dentro de las trayectorias. Para esto, se emplea el principio de Descripción de Longitud Mínima (MDL), que busca minimizar la cantidad de información requerida para representar las trayectorias. Este enfoque permite filtrar los puntos que contribuyen de manera significativa a la forma general de la trayectoria, eliminando aquellos que son redundantes o que aportan poca información.

La selección de puntos significativos se realiza analizando cada trayectoria y determinando aquellos puntos que, al ser eliminados, resultarían en la mayor pérdida de información según el criterio MDL. Este proceso resulta en una versión simplificada de la trayectoria original, compuesta solo por sus puntos más representativos.

**Implementación:** Se analiza cada trayectoria para determinar puntos donde su eliminación aumentaría significativamente el error de representación según el criterio MDL. La trayectoria se segmenta en estos puntos, resultando en una serie de segmentos de línea que capturan la esencia del movimiento original con una menor cantidad de datos.

**3. Agrupamiento de Trayectorias**

Con los puntos significativos identificados, el siguiente paso es agrupar las trayectorias simplificadas en clusters. Para esto, se utiliza un algoritmo de clustering basado en densidad, como DBSCAN, que agrupa las trayectorias según su proximidad espacial y la densidad de puntos significativos. Este método considera como miembros del mismo cluster a aquellas trayectorias que están suficientemente cerca una de otra según la función de distancia compuesta definida anteriormente.

**Implementación:** Se evalúa cada segmento de línea para determinar si cumple con los criterios de un segmento núcleo, basándose en el número de vecinos dentro de un radio especificado (ε) y un número mínimo de segmentos de línea (MinLns). A partir de los segmentos núcleo, se expanden los clusters agregando segmentos directamente alcanzables, formando grupos de segmentos con características de movimiento similares.

**4. Representación de los Clusters**

Finalmente, una vez formados los clusters, se procede a generar una representación de cada uno. Esto implica identificar y destacar las trayectorias representativas dentro de cada cluster, que sirven como ejemplos claros de las características comunes del grupo. Estas trayectorias representativas se seleccionan basándose en su centralidad dentro del cluster y su similitud con otras trayectorias del mismo grupo.

**Implementación:** Se calcula una trayectoria representativa para cada cluster determinando la dirección promedio de los segmentos de línea y luego promediando las posiciones de los segmentos a lo largo de esta dirección. Este proceso resulta en una serie de puntos que capturan el patrón de movimiento predominante de los segmentos dentro del cluster.