

R en movimiento: una revisión de los paquetes para analizar movimiento dirigida a usuarios y desarrolladores

Abstract

El creciente avance de la tecnología de seguimiento o tracking (e.g. GPS, cámaras de video, acelerómetros), está dando paso a grandes cantidades de datos que permiten monitorear el movimiento tanto de seres humanos como de animales. En paralelo, se van desarrollando variadas y sofisticadas herramientas para procesar, visualizar y analizar datos de seguimiento. Sólo en R, hemos encontrado 59 paquetes focalizados en estas tareas (e.g. `adehabitatLT`, `move`).

En esta presentación, revisamos y describimos brevemente los 59 paquetes, que llamamos paquetes de seguimiento, basándonos en un flujo de trabajo centrado en este tipo de datos y dividido en tres etapas: pre-procesamiento, post-procesamiento y análisis. El análisis lo dividimos a su vez en visualización de datos, descripción de trayectorias, reconstrucción de trayectorias, identificación de patrones comportamentales, caracterización del uso del espacio, simulación de trayectorias y otros.

Utilizando un análisis de redes, evaluamos las relaciones entre los paquetes y mostramos que un tercio de ellos trabaja en aislamiento. Este resultado es un reflejo de una fragmentación en la comunidad de programadores del movimiento en R (Fig. 1).

A la luz de los resultados encontrados, damos algunos criterios para potenciales usuarios para escoger paquetes, y sobre todo, criterios para desarrolladores, con el propósito de maximizar la utilidad de su paquete y fortalecer los vínculos dentro de la comunidad programadora. Si bien estos criterios están inspirados en los paquetes de seguimiento, pueden ser extrapolados a todo tipo de paquete en R.

El manuscrito relacionado a este trabajo se encuentra en: <https://arxiv.org/abs/1901.05935>; los datos sobre cada paquete se encuentran en <https://doi.org/10.5281/zenodo.3066226>.

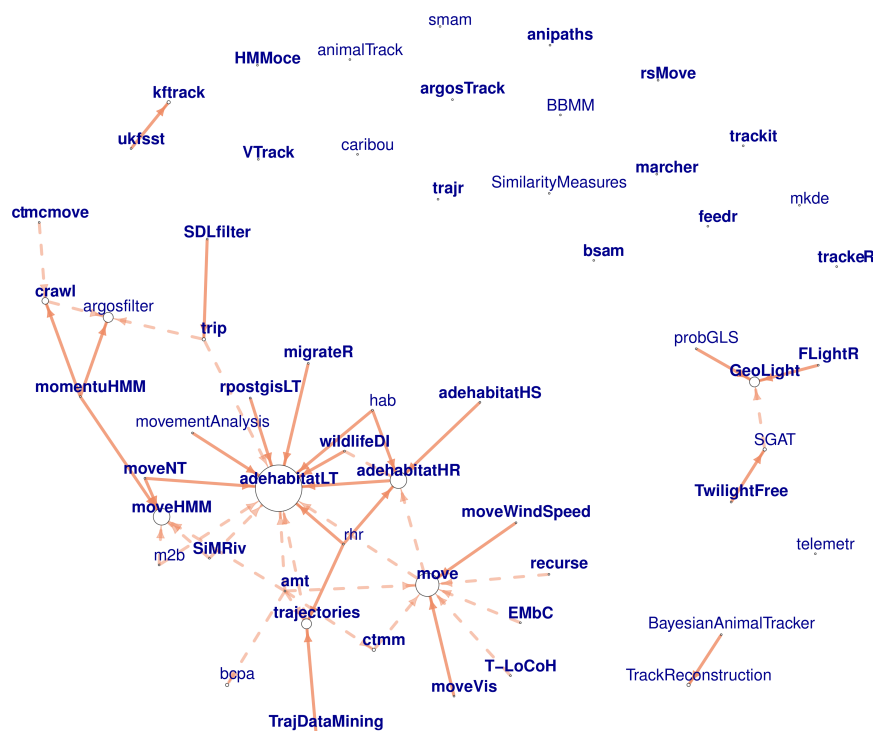


Figura 1: Representación de la red de dependencias y sugerencias entre paquetes de seguimiento. Las flechas van hacia el paquete que los otros sugieren (flechas discontinuas) o dependen de (flechas sólidas). Las fuentes en negrita corresponden a paquetes activos (última actualización hecha hace menos de 1 año). El tamaño del círculo es proporcional al número de paquetes que sugieren o dependen de éste.

Consideramos que este trabajo puede ser un punto de partida para discutir, desde América Latina, la relevancia de nuestras propuestas de innovación y desarrollo de la mano de R.

Keywords: biologging, R packages, movement ecology, tracking data, trajectometry