PhD Project Show Case: Insight into Animal Movement Data Using Machine Learning Techniques

Maryam Teimouri and Håvard Tveite

Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences (NMBU), N-1432 Norway

maryam.teimouri@nmbu.no, havard.tveite@nmbu.no

Introduction

The emergence of Global Navigation Satellite System (GNSS) has simplified the acquisition of spatio-temporal information of moving objects including animals. In this research project, we focus on some of the concerns of animal scientists from Norwegian University of Life Sciences (NMBU) about movements of freely roaming livestock, i.e. sheep in our case, equipped with GPS collars, when being released on open ranges. Understanding of how and where sheep move and discovering of their movement patterns is worthwhile in order to better allocate resources and/or assess the impact of utilization by livestock. For instance, detecting presence of predator based on the extracted movement patterns and changing environmental conditions from anomalous sheep behavior could prevent the loss of livestock. The research project (“Cyber Shepherd”) mainly focuses on the following main areas: (1) monitoring activities of sheep in order to detect interesting behavior, (2) classification and identification of movement patterns from trajectory data, (3) creating models of animal movements, particularly sheep, (4) investigating of requirements for spatial and temporal resolution of trajectory data, (5) utilizing digital terrain models in order to improve the accuracy of observations, and (6) real-time monitoring and analysis of movement data.

Test Data and Case Location

We have started our research based on a test data collected on a pasture, located in Hol municipality, Buskerud country in Norway, including 9 enclosures called A, B, C, D, E, F, G, H, and I, with overall area of 2.7 Km². Ewes were tracked from July 1st to September 13th, 2012, within 6 enclosures (Table 1) with temporal granularity of 30 minutes.

Preliminary Analyses

At the first phase, we have used some of the existing exploratory data analysis techniques in order to comprehend and summarize the characteristics of test data without any priori hypothesis. The boxplots of elevation and step-length variation for monitored ewes are almost the same as Figure 1. For example, we have concluded that sheep in the pasture were less active in the high elevation during night and more active in the low elevation during day.
In the second phase, we have extracted those areas where frequently visited by sheep. We have applied DBSCAN\(^1\) algorithm (Ester et al. 1996) to the test data (Figure 2). As we expected salt locations and water sources are of interesting area where frequently visited by sheep.

---

\(^1\) Density-Based Spatial Clustering of Application with Noise

---

<table>
<thead>
<tr>
<th>Enclosure</th>
<th>Area (m(^2))</th>
<th>Number of Sheep</th>
<th>Area (m(^2)) per Sheep</th>
<th>Sheep per Area (m(^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group B</td>
<td>314385,0</td>
<td>18</td>
<td>1,75E+04</td>
<td>5,73E-05</td>
</tr>
<tr>
<td>Group C</td>
<td>339037,9</td>
<td>6</td>
<td>5,65E+04</td>
<td>1,77E-05</td>
</tr>
<tr>
<td>Group D</td>
<td>383975,0</td>
<td>8</td>
<td>4,80E+04</td>
<td>2,08E-05</td>
</tr>
<tr>
<td>Group F</td>
<td>220673,0</td>
<td>17</td>
<td>1,30E+04</td>
<td>7,70E-05</td>
</tr>
<tr>
<td>Group H</td>
<td>241938,5</td>
<td>17</td>
<td>1,42E+04</td>
<td>7,03E-05</td>
</tr>
<tr>
<td>Group I</td>
<td>230959,0</td>
<td>5</td>
<td>4,62E+04</td>
<td>2,16E-05</td>
</tr>
</tbody>
</table>

Table 1. Dataset overview

---

Figure 1. Exploratory data analysis (e.g. elevation and step-length variation)
Trajectory is considered as a polyline of consecutive points ordered in time. Different static and dynamic behavior of sheep can be detected from such time-ordered locations such as grazing and dozing. We have accomplished some data pre-processing methods to clean the test data of possible outliers. For example, the captured points that have a distance from the previous points of more than three times the standard deviation of the distances between consecutive captured points were eliminated (Dodge et al. 2009). Then, we have segmented trajectories at where the time lag between two consecutive location records was greater than 1.5 hours, i.e. 3 missing locations while one or two missing location records were resampled by linear interpolation. Now, Based on Adehabitat Package (Calenge 2006), we have regular trajectories that are characterized by constant time lag between successive locations. Descriptive parameters such as turning angle, sinuosity as movement features of trajectories may be derived to feed into the machine learning algorithm in the future analyses (for example see (Godsk and Kjærgaard 2011; Dodge et al. 2009).

As the next step, we have identified the locations where the sheep standing-still (i.e. stops). As we know even when sheep stays at one location for a while, due to error related to GPS coordinates, we do not have the same coordinates for that time interval. We were interested in discovering those consecutive locations actually were at the same locations but in our data looks different. For this purpose, we have applied the CB-SMoT as a Clustering Based technique for detecting Stops and Moves of Trajectories (Palma et al. 2008) (Figure 3). CB-SMoT is the adapted version of DBSCAN looking for clusters of points along a trajectory; therefore notion of neighborhood is confined to points on the considered trajectory inside a temporal window. The circular histograms summarize the times of extracted stop points. It is observed from Figure 3, majority of stop points happened around midnights when sheep sleep. Manipulating input parameters of CB-SMoT algorithm can result in discovering the other stop points happened during middays.
Future work

Next, we are going to find out the relationships between the constructed trajectories, discover similar sub-trajectories, and cluster different behavior groups from the tracked sheep.

Our plan is to design a new field experiment by using new equipment such as camera and activity sensors with a finer temporal resolution.

Attending to the GIScience 2014 Workshop on Analysis of Movement Data (AMD'14) opens many doors to me to improve my understanding in the aforementioned issues.

References


