動物の移動行動クラスタリングとセンサーネットワークシミュレーション

川瀬純也*

Spatiotemporal Movement Clustering and Sensor Network Simulation of Animals

Junya Kawase*

Abstract: This study demonstrates the potential of animal wearable sensor networks and aims to efficiently expand their coverage area by evaluating the data collection performance of delay-tolerant networking (DTN) through simulations based on actual animal movement data. GPS trajectories of dairy cows were collected over 111 consecutive days and analyzed using spatiotemporal movement clustering. Based on the clustering results, simulations were conducted assuming data delivery and collection via DTN, and the coverage area and coverage ratio of the sensor network were compared. In scenarios where individuals were selected across clusters, the results were largely consistent with those of biased selection scenarios, but the risk of complete data collection failure was reduced.

Keywords: 移動行動クラスタリング（spatiotemporal movement clustering），カバー面積（coverage area），動物装着型センサー（animal wearable sensor）

１．はじめに

野生動物に様々なセンサーデバイスを携行してもらうことで，人間では容易に侵入することができない自然環境の情報や，野生動物同士の接触履歴を知ることができ，社会の様々な場面での活用が期待できる．森林や山地に富んだ日本の自然環境下では，記録したデータを動物自らが運搬し，他のデバイス装着個体と遭遇した際にそのデータを共有し，これを繰り返すことで数珠つなぎにデータを伝播・回収する“遅延耐性ネットワーク”（DTN）という技術の利用が想定されている（Kobayashi et al., 2018）（図 1）．

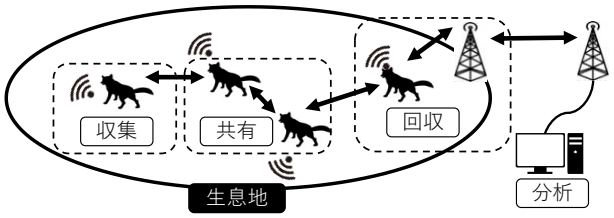


図 1　野生動物 DTN のコンセプト

限られたデバイス装着個体数で，効率的にセンサーネットワークがカバーする面積を広くするためには，異なる群れの間を行き来するような個体や，他の個体に接触する機会が多いような個体にセンサーデバイスを携行してもらうことが理想的だと考えられる．しかし，デバイス装着個体の選択やそれによって変化するセンサーネットワークのカバー面積について，実際の動物の移動行動データを用いた実験や評価を行った先行研究はない．

そこで本研究では，自由に昼夜放牧されている家畜の GPS データを用いて，その移動行動をクラスタリングするとともに，各クラスター間を行き来するような個体を効率的に選択する手法について検討を行い，実際に DTN が構築されていた場合を想定してセンサーネットワークがカバーする面積について評価を行った．

２．使用した GPS データ

本研究では，北海道のある牧場において，夏季期間，自由に昼夜放牧されている雌の乳牛の GPS データを用いた．放牧地の概略図を図 2 に示す．放牧地の広さは約 40 ha，高低差は最大で約 150 m．東側の低い場所に牛舎が置かれている．

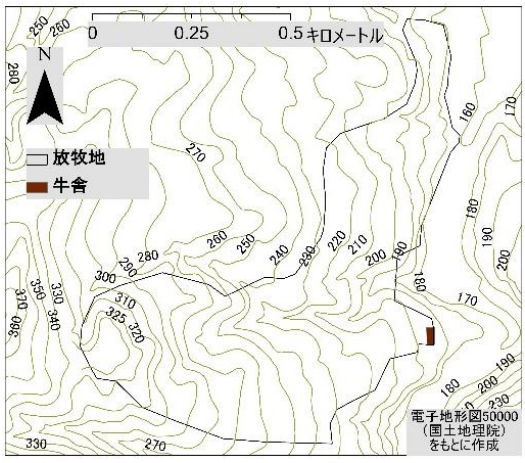


図 2　放牧地の概略図

GPS ログは牛の首に装着したロガーにより 20 分ごとに記録され，無線通信で収集される．牛の体勢や樹木等

*　正会員　　東京大学情報基盤センター（Information Technology Center, the University of Tokyo）
〒277-0882　千葉県柏市柏の葉 6-2-3　E-mail：j.kawase@ds.itc.u-tokyo.ac.jp

の影響により GPS ログが記録・収集できなかった場合は，そのまま欠損として扱った．使用したロガーは GlobalSat 社の LT-520 で，牛用の首輪を用いて装着した．GPS の測位タイミングは，ロガーの設定や首輪の取り付けとメンテナンスの影響を受け，全てのロガーで一致させることはできなかった．本研究では，2023 年 6 月 16 日から 10 月 4 日の連続した 111 日間において，データを十分に回収できた 21 頭の雌の乳牛を対象とした．

牛たちは午前 6 時ごろと午後 3 時ごろの 1 日 2 回，搾乳や管理のため，放牧地から牛舎に自ら戻るように馴致されており，まれに戻らない個体がいる場合は，農家の人によって誘導される場合もある．また放牧地には雄の乳牛が 2 頭放牧されており，行動を共にしている．

3．移動行動クラスタリング手法の概要

3.1　配列アライメントを用いた移動行動クラスタリング手法の関連研究

本研究では，移動行動のクラスタリングには配列アライメント手法を用いた．もともとはバイオインフォマティクスにおいて，DNA などの配列の類似した領域を特定する際などに用いられる手法であるが，時空間行動分析と近い分野での例として，日常行動を対象とした日記調査の分析への援用が挙げられる（Wilson, 1998; Wilson et al., 1999）．観光者の GPS データを対象とした例（Shoval and Isaacson, 2007），動物園などの限られた空間での観覧者を対象とした例（矢部, 2010; 川瀬・伊藤, 2016）などがある．動物の移動行動に同様の手法を使用した例は見られないが，類似した手法として，野生動物の移動行動データを文字列化して最長共通部分列（LCSS）を用いてクラスタリングを行い，動物種間の違いを明らかにした研究などがある（Gao et al., 2016）．いずれにしても，毎日繰り返され，しかし同じではない移動行動を，一定期間ごとに区切ってクラスタリングを行い，さらに期間全体でのクラスタリングを行い，活用しようというような例はみられない．

3.2　移動行動の文字列化

午前 6 時からの 24 時間を 1 日のサイクルと捉え，20 分ごとに滞在したエリアを表す文字を並べて，1 日の移動行動を文字列として表す．ある 20 分間の GPS ログが欠損している場合は空白（indel）として扱った．

人間の移動行動を対象とした分析の場合は，対象エリアの形状や機能をもとにエリア分けがなされるが，放牧地では同様の手法をとることができない．そこで，多くの GPS ログが記録された各々の地点が，その日の乳牛の行動の特徴を表していると仮定し，その日の全ての GPS ログをプロットしてカーネル密度推定を行い，密度の頂点から作成したボロノイ図を文字列化のためのエリア分けとした（図 3）．エリア分けは 111 日間それぞれで作成した．20 分ごとに各 GPS ログが位置するエリアを表す文字あるいは空白を表す indel を順番に並べてゆき，24 時間当たり 72 文字の文字列でもって 1 日の移動行動を表した．36 文字以上が indel となる場合は，その個体の 1 日全体を欠損として扱った．

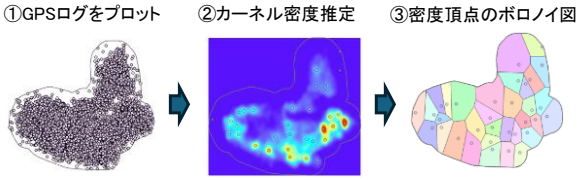


図 3　エリア分けの作成手順イメージ

3.3　配列アライメントの実行

配列アライメントの計算には R パッケージ TraMineR（Gabadinho et al., 2011）を用いた．ある文字列を別の文字列に変換する際の編集操作のコストを計算する Levenshtein 距離の概念をもとにした Optimal matching のアルゴリズムを用いた．エリア分けをもとに，同じ文字の場合は「2」，隣接するエリア同士の置換の場合は「1」，indel との置換の場合は「1」，それ以外の置換の場合を「0」とし，文字列間の類似度を計算した．最後に階層的クラスタリングを行い，樹状図を得た．

1 日ごとにクラスタリングを行い，さらにそのクラスタリング結果 111 日分を並べて文字列とし，111 日分の期間全体でのクラスタリングを行った．1 日全体が欠損となった個体は，当該日の 1 日ごとのクラスタリングの対象からは除外し，期間全体でのクラスタリングでは当該日のクラスタリング結果を indel として扱った．

4．クラスタリング結果

4.1　1 日ごとのクラスタリング結果の例

ある 1 日のクラスタリング結果によって得られた樹状図を図 4 に，クラスターごとに分けて描画した時空間パスを図 5 に示す．この時空間パスでは，地表面を午前 6 時とし，1 分ごとに 1 m の高さを取るように移動軌跡を描画している．この例では，特に 1 日の前半と後半の違いに着目することで，クラスターごとに移動行動の特徴の違いを見て取ることができる．

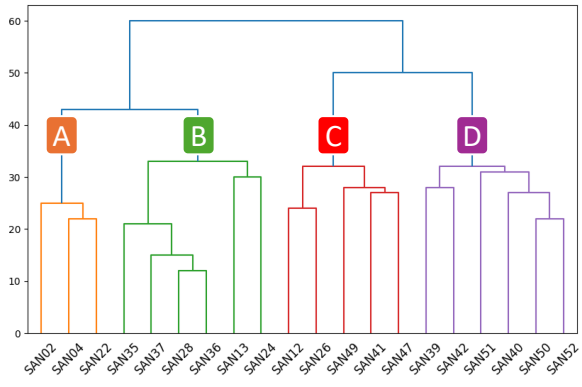


図 4　クラスタリング結果の樹状図の例

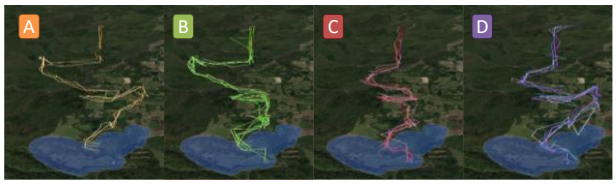


図 5　クラスタリング結果ごとの時空間パスの例

4.2　期間全体でのクラスタリング結果

111 日間の期間全体でのクラスタリング結果の樹状図を図 6 に示す．図中，左端の 2 個体は 1 日全体での欠損が多く含まれていることがわかった．図 7 に 111 日間全体での文字列の一部を示す．1 行ごとが各個体を表し，1 日ごとのクラスタリング結果が左から右に順に並んでいる．1 列が 1 日ごとの結果を表しており，期間全体での結果を参考に，その日のクラスターごとに彩色している．この結果から，多くの日で移動行動を共にしているクラスターを得ることができたことが示唆された．

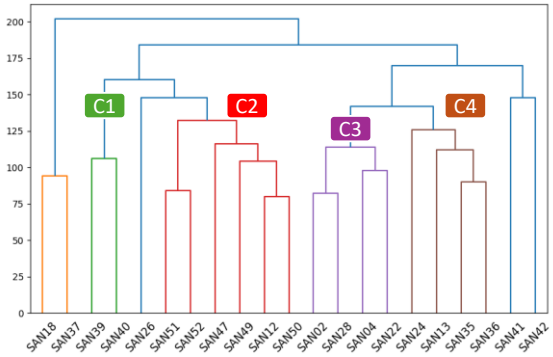


図 6　期間全体でのクラスタリング結果の樹状図

ID	全体	6/16	6/17	6/18	6/19	6/20	6/21	6/22	6/23	6/24	6/25	6/26	6/27	6/28	6/29	6/30	7/1
SAN18																	
SAN37		c1	c1	c2	c1	c3	c3	c3	c2	c1	c2	c2	c1	c1	c2	c2	c2
SAN39	C1	c3	c2	c1	c2	c3	c3	c3	c1	c3	c2	c2	c2	c1	c2	c1	c4
SAN40	C1	c1	c2	c2	c2	c3	c5	c2	c1	c2	c2	c2	c2	c1	c3	c1	c4
SAN26		c2	c1	c1	c1	c2	c0	c3	c2	c0	c1	c2	c2	c1	c2	c2	c3
SAN51	C2	c1	c3	c2	c1	c3	c0	c3	c1	c3	c2	c2	c1	c1	c4	c2	c4
SAN52	C2	c3	c3	c1		c3	c1	c3	c1	c1	c1	c2	c1	c1	c4	c2	c4
SAN47	C2	c3	c1	c2	c2	c3	c5	c3	c2	c1	c2	c1	c2	c1	c2	c1	c3
SAN49	C2	c3	c1	c2	c2	c2	c5	c2	c2	c1	c1	c2	c2	c0	c4	c2	c3
SAN12	C2	c3	c3	c1	c2	c3	c2	c3	c2	c1	c2	c2	c2	c1	c4	c2	c3
SAN50	C2	c3	c3	c1	c2	c2	c5	c3	c2	c1	c1	c2	c1	c1	c1	c2	c4
SAN02	C3	c1	c2	c2	c1	c1	c3	c1	c2	c3	c2	c3	c1	c1	c2	c1	c1
SAN28	C3	c1	c2	c2	c1	c1	c4	c3	c2	c3	c2	c3	c1	c1	c2	c1	c2
SAN04	C3	c1	c2	c2	c1	c1	c4	c1	c2	c3	c2	c3	c1	c1	c4	c1	c1
SAN22	C3	c1	c2	c2	c1	c1	c4	c1	c2	c3	c2	c1	c2	c1	c1	c1	c1
SAN24	C4	c1	c2	c2	c1	c1	c2	c1	c2	c2	c2	c3	c1	c1	c2	c1	c2
SAN13	C4	c1	c3	c2	c1	c1	c3	c1	c2	c1	c2	c3	c2	c1	c2	c2	c2
SAN35	C4	c1	c1	c2	c1	c1	c1	c1	c2	c1	c2	c3	c2	c1	c4	c1	c2
SAN36	C4	c1	c1	c2	c1	c1	c3	c1	c2	c3	c2	c3	c2	c1	c2	c1	c2
SAN41		c1	c1	c2	c2	c1	c0	c3	c2	c1	c2	c1	c1	c1	c2	c1	c3
SAN42		c2	c3	c1	c2	c3	c5	c2	c0	c1	c2	c2	c2	c1	c3	c2	c4

図 7　111 日間のクラスタリング結果を並べた文字列の一部

5．DTN がカバーする面積のシミュレーション

実際の GPS データをもとに，乳牛たちに装着したデバイスによって DTN 型のネットワークが構成されている場合を想定し，センシング，データの伝播・回収のシミュレーションを行い，センサーネットワークがカバーする面積を推定する．一部の個体のみで DTN が構成される場合を想定し，個体の組み合わせによってセンサーネットワークがカバーする面積がどのように変化するかを評価する．4 章の移動行動クラスタリングにより得られた各クラスターを参考に，DTN を構成する個体を選択した場合と，それと反するように個体を選択した場合との結果を比較した．

5.1　シミュレーションの設定

本研究では，シミュレーションは 1 分単位で進行し，1 日単位（午前 6 時から翌朝午前 5 時 59 分）で評価するものとした．GPS ログの欠損部分は，存在する GPS ログ間を直線でつなぎ，ログ間の時間差で直線を等分することで，1 分単位で補間した．これは，収集された GPS ログの測位タイミングが必ずしも一致しないことへの対応も含んでいる．

各デバイスは平面距離で 30 m 以内に 1 分以上継続して位置する場合に通信が可能であるとした．なお，データ量にかかわらず通信が可能であればデータの送受信が完了するものとした．1 日に 2 度，基本的に全ての個体が牛舎に集まることを考慮し，牛舎があるエリアから平面距離で 100 m の範囲では通信が行えないものと設定した．

5.2　シミュレーションの進行

①シミュレーションは 1 分 1 ターンごとに進行し，通信範囲内に位置する個体を抽出する．②1 つ前のターンから継続して通信範囲内に位置した個体同士は通信が成功したとみなし，それまでに収集した観測データを共有する．③次のターンに遷移する前に，フィールドでの観測が行われ，観測データ（本研究では，デバイスの ID，時刻，座標のみ）が追加される．④これを繰り返し，最終ターンまでに固定ノードに到達した観測データを「回収したデータ」とする．なお，固定ノードは実際にロガー回収用のゲートウェイが設置されている位置とし，通信可能範囲は 30 m に設定した．

5.3　「回収したデータ」の評価

放牧地全体を 25 m × 25 m のグリッド（図 8）でわけ，「回収したデータ」が位置するグリッドを「回収できたグリッド」とみなす．このグリッドを集計し，各シナリオを以下の項目で評価する．なお，GPS ログがひとつも記録されていないグリッドは評価対象の全体数から除外した．

1. 回収できたグリッドの数
2. 全個体を対象とした DTN によって回収できたグリッド数を 100 とした場合のカバー率

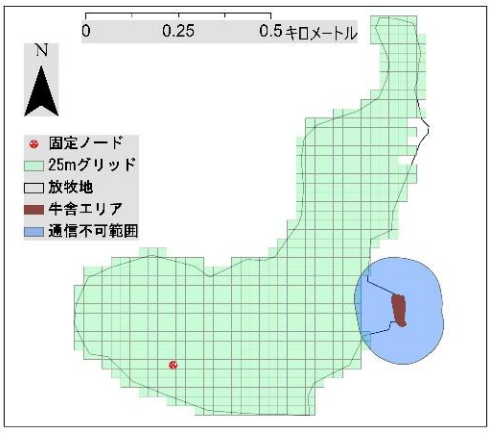


図 8　評価用 25 m × 25 m グリッド

5.4　DTN 構成個体選択シナリオの設定

21 個体のうち，欠損の多い 2 個体を除いた 19 個体を対象とすることとした．なお，19 個体のみで 4 章と同様のクラスタリングを再度行ったが，得られる結果は変わらなかった．次に，111 日間のうち，19 個体のいずれかでも 1 日全体の欠損が含まれる日は評価の対象外とした．そして，19 個体のいずれもが固定ノードの 30 m 以内に接近せずに，データを回収できなかった日も評価の対象外とし，111 日間のうち 62 日分を評価の対象とした．

クラスタリング結果から，C1（2 個体），C2（6 個体），C3（4 個体），C4（4 個体）からそれぞれ 1 個体ずつ選び，残る 3 個体を加えた全 7 個体で構成する DTN を想定し，192 パターンでシミュレーションを行った（シナリオ A）．比較対象として，似通った移動行動の個体たちに偏ってしまった場合を想定し，C2・C3・C4 のうち 2 つのクラスターから 7 個体を選択した 248 パターンのシミュレーションを行った（シナリオ B）．

5.5　シミュレーション結果

シミュレーション結果の集計結果を箱ひげ図で図 9 に示す．回収できたグリッド数およびカバー率のどちらについても，最大値，平均値，中央値，四分位範囲などでシナリオ A が上回ってはいるがその差はわずかであった．

各シナリオにおいて，回収できたグリッドが 0 だったパターンの数を表 1 に示す．シナリオ B の方が，回収できたグリッド数が 0 であったパターンが含まれる日数が明らかに多い．各クラスターから個体を選択したシナリオ A の方が，固定ノードまでデータを運んでくれる個体が含まれる確率が高くなり，データ回収が失敗するリスクを低減できる可能性が示された．

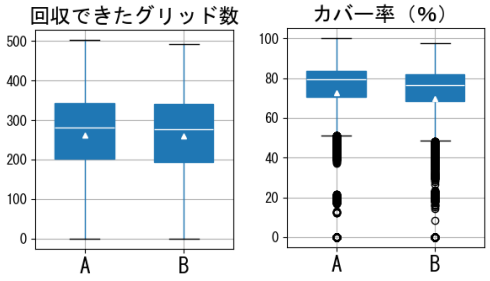


図 9　シナリオ A および B の結果比較

表 1　回収できたグリッドが 0 だったパターンの数

	6月				7月							
	18日	24日	29日	30日	5日	10日	13日	16日	17日	22日	23日	31日
シナリオA		6		16	160		144		24			160
シナリオB	1		1		80	1	157	1	9	248	248	80

	8月							9月				計
	1日	2日	17日	20日	26日	29日	31日	12日	17日	18日	20日	
シナリオA					144			32			45	731
シナリオB	1	8	8	16	157	9	1	8	1	8	16	1059

おわりに

本研究では，野生動物装着型 DTN がカバーする面積を効率的に広くすることを目的に，自由に昼夜放牧されている家畜の GPS データを用いて，その移動行動をクラスタリングし，DTN を構成する個体の選び方でカバーする面積がどのように変化するのかをシミュレーションして評価を行った．

提案手法において各クラスターから個体を選択するだけでは，カバーする面積を効率的に広くできるとは言えなかった．しかし，データ回収が失敗するリスクを低減できる可能性が示された．

今後は，期間全体でのクラスタリング方法の再検討やシミュレーションの評価方法の再検討を行っていくとともに，現実的な野生動物 DTN の運用方法に沿ったセンサーデバイス装着個体選択方法を検討していく．

謝辞

本研究は，JSPS 科研費 JP23K12323 の助成を受けたものです．また本研究成果の一部は，データ活用社会創成プラットフォーム mdx を利用して得られたものです．

参考文献

川瀬純也・伊藤史子（2016）対象者の状態を加味した配列アライメントによる時空間行動分析手法の提案と分析 - 歩行速度と観覧対象との距離を活用した観覧行動類型化手法の提案 - .「地理情報システム学会講演論文集」，**25**，E-6-1.

矢部直人（2010）GPS データに対する配列解析の援用.「地理情報システム学会講演論文集」，**19**，181-190.

Gabadinho, A., Ritschard, G., Müller, N.S. & Studer, M. (2011) Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software*, **40**(4), 1-37.

Gao, P., Kupfer, J. A., Zhu, X., & Guo, D. (2016) Quantifying animal trajectories using spatial aggregation and sequence analysis: a case study of differentiating trajectories of multiple species. *Geographical Analysis*, **48**(3), 275-291.

Kobayashi, H. H., Nakagawa, K., Makiyama, K., Sasaki, Y., Kudo, H., Niraula, B., & Sezaki, K. (2018) Animal-to-animal data sharing mechanism for wildlife monitoring in Fukushima exclusion zone. *Multimodal Technologies and Interaction*, **2**(3), 40.

Shoval, N., & Isaacson, M. (2007) Sequence alignment as a method for human activity analysis in space and time. *Annals of the Association of American geographers*, **97**(2), 282-297.

Wilson, W. C. (1998) Activity pattern analysis by means of sequence-alignment methods. *Environment and Planning A*, **30**(6), 1017-1038.

Wilson, C., Harvey, A., & Thompson, J. (1999) ClustalG: Software for analysis of activities and sequential events. *IATUR conference proceedings*.