Inference of Leadership of Coordinated Activity in Time Series

by

CHAINARONG AMORNBUNCHORNVEJ
B.Eng., King Mongkut's Institute of Technology Ladkrabang,Thailand, 2011
M.Eng., King Mongkut's Institute of Technology Ladkrabang,Thailand, 2013

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Computer Science
in the Graduate College of the
University of Illinois at Chicago, 2018

Chicago, Illinois

Defense Committee:
Tanya Y. Berger-Wolf, Chair and Advisor
Elena Zheleva
Brian Ziebart
Gyorgy Turan, Mathematics, Statistics, and Computer Science
Eamonn Keogh, University of California, Riverside

ACKNOWLEDGMENTS

The thesis has been complete with the support from many people. I would like to take this occasion to express my appreciation toward all of them. First of all, thank you my advisor, Prof. Tanya Berger-Wolf, who always patiently guides me since the day one when I had no background to mathematically prove any claiming statement until I gained enough skills to be an independent researcher. She always encourages me to have a freedom to thrive my ideas. I am not afraid to encounter any research challenges and failures because of her support. I am very grateful to all members of my thesis committee, Prof. Keogh, Prof. Turan, Prof. Zheleva, and Prof. Ziebart, who gave insightful feedback and suggestions. I also would like to thank all professors and staff as UIC who provided supportive resource for me to guarantee my success. Especially, thank you Vandana Loebel and all ITA staff who provided the support for my English language development. Without them, I cannot survive even my first semester. I also would like to thank my lab mates, Ivan Brugere, Vena Jia Li, Aynaz Taheri, and the rest of Comp Bio alumni, as well as my co-authors, Dr. Strandburg-Peshkin, Dr. Farine, and Prof. Crofoot, for their helpful supports. I am grateful to the precious support from my love, Usawadee Chaiprom, and my family who are always beside me no matter what happens. Finally, this thesis is dedicated to everyone in my life, who always supports and encourages me.

CA

CONTRIBUTION OF AUTHORS

In this dissertation, it consists of the works from three previously published papers: the works in [1–3] (Chapter 5), [4] (Chapter 6), and [5] (Chapter 7). I would like to take this opportunity to give credits to my fantastic co-authors. I really appreciate their contributions to my success. All works were under the supervision of Prof. Berger-Wolf. For the work in [1–3], I was the main author who conducted both experiments and analysis. Mr. Brugere provided us the NASDAQ stock-market dataset as well as edited and analyzed many sections in the paper. Dr. Strandburg-Peshkin, Dr. Farine, and Prof. Crofoot edited the introduction section in the biology part as well as provided insightful suggestions. For the work in [4], I mainly conducted the analysis and experiments under insightful suggestions and editing effort by Prof. Berger-Wolf. For the work in [5], Prof. Crofoot and Prof. Berger-wolf helped me editing the paper as well as providing valuable recommendations.

Thank you,

Chainarong Amornbunchornvej

TABLE OF CONTENTS

CHAPTER		PAGE
1	**INTRODUCTION**	1
1.1	Motivation	1
1.2	Related work	2
1.2.1	Computational leadership categories	3
1.2.1.1	Movement leadership	3
1.2.1.2	Influence maximization	4
1.2.2	Coordination mechanisms inference from time series	5
1.2.3	Mining and modeling frequent patterns of leadership dynamics	7
1.2.4	Limitations of current methods	8
1.3	Our proposed solutions	11
2	**PRELIMINARIES**	13
2.1	Coordination event detection and initiator identification in time series	13
2.1.1	Influence maximization vs. coordination initiator inference problem	16
2.2	Inferring leadership dynamics of complex movement from time series	17
2.3	Identifying traits of leaders in movement initiation	18
2.4	Coordination mechanisms inference from time series	20
2.5	Mining and modeling complex leadership dynamics of movement data	21
3	**PROBLEM FORMALIZATION AND ANALYSIS**	24
3.1	Single coordination event without noise	24
3.2	Useful observations	27
3.2.1	Following relation with noise	29
3.3	Multiple coordination events without noise	31
3.4	Multiple coordination events with noise	33
3.4.1	Coordination measure	33
3.5	Coordination mechanism inference	35
3.5.1	Problem formalization	37
3.6	Mining patterns of leadership dynamics	38
4	**LEADERSHIP MODELS AND DATASETS**	40
4.1	Simulation of leadership models	40
4.1.1	Dictatorship model (DM)	40

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

4.1.2 Hierarchical model (HM) . . . 41
4.1.3 Event-based model (EM) . . . 42
4.1.4 Initiator model (INIT-k) . . . 42
4.1.5 Crowd model (CM) . . . 43
4.1.6 Linear Threshold model (LT) . . . 44
4.1.7 Independent Cascade model (IC) . . . 45
4.1.8 Random model . . . 46
4.2 Real-world datasets . . . 47
4.2.1 Baboon trajectories . . . 47
4.2.2 Fish schools trajectories . . . 47
4.2.3 Stock closing-price time series . . . 47

5 FLICA: A FRAMEWORK FOR LEADER IDENTIFICATION IN COORDINATED ACTIVITY . . . 49
5.1 Introduction . . . 49
5.1.1 A working example . . . 50
5.2 Methods . . . 52
5.2.1 Following relation inference . . . 52
5.2.2 Dynamic following network inference . . . 54
5.2.3 Coordination intervals detection . . . 55
5.2.4 Ranking comparison . . . 57
5.2.4.1 PageRank . . . 57
5.2.4.2 Velocity Convex Hull . . . 58
5.2.4.3 Position Convex Hull . . . 59
5.2.5 Leadership model features . . . 60
5.2.6 Local vs. Global Matching . . . 62
5.3 Experimental setup . . . 63
5.3.1 Synthetic trajectory simulation . . . 63
5.3.2 Evaluation . . . 63
5.3.3 Compared leadership methods . . . 64
5.3.4 Sensitivity analysis . . . 66
5.3.4.1 Support of faction leading . . . 67
5.3.4.2 Simulation and accuracy measure . . . 69
5.3.4.3 Position and Direction noises . . . 70
5.3.4.4 Time window sensitivity . . . 70
5.4 Results . . . 70
5.4.1 Identifying leaders . . . 70
5.4.2 Case study: trained initiators in fish schools . . . 72
5.4.3 Case study: finding “initiators” of stock market events . . . 72
5.4.4 Case study: baboon activity classification by following network density . . . 76
5.4.5 Leadership model classification . . . 79

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

5.4.5.1 Baboon leadership model characterization 81
5.4.6 Following network density perturbation in baboon and fish data 81
5.4.7 Sensitivity analysis 84
5.5 Discussion 88

6 MFLICA: A FRAMEWORK FOR MULTIPLE-FACTIONS LEADER IDENTIFICATION IN COORDINATED ACTIVITY 90
6.1 Introduction 90
6.2 Methods 92
6.2.1 Factions detection and coordination intervals. 92
6.2.2 Time window inference. 93
6.2.3 Leadership comparison. 94
6.3 Evaluation Datasets 94
6.3.1 Leadership models. 94
6.3.2 Synthetic trajectory simulation. 94
6.3.2.1 Linear coordination event. 95
6.3.2.2 Splitting/Merging coordination event. 96
6.3.3 Biological datasets 97
6.3.3.1 Baboon trajectories. 97
6.3.3.2 Fish schools trajectories. 97
6.4 Evaluation criteria 97
6.4.1 Individual assignment. 97
6.4.2 Leadership prediction. 97
6.4.3 mFLICA Time complexity 98
6.5 Results 98
6.5.1 Leadership Identification. 98
6.5.2 Case study: trained leaders in fish schools. 101
6.5.3 Case study: detecting the group merging event of baboons. . 102
6.5.4 Centrality measures in multi-faction datasets 102
6.6 Discussion 106

7 TRAITS OF LEADERS FRAMEWORK 108
7.1 The Proposed Approach 108
7.1.1 Bidirectional agreement in multi-agent systems 108
7.1.2 Bidirectional Agreement condition 110
7.1.3 Leaders as state changers 111
7.1.4 Approach overview 112
7.1.5 Leadership trait characterization scheme 113
7.1.5.1 The quantification of the traits of interests. 113
7.1.5.2 Convex hull ranking measures. 114
7.1.5.3 Rank correlation. 116
7.1.5.4 Leadership model classification using trait rank correlation . . 119

CHAPTER **PAGE**

7.2 Experimental setup . . . 119
7.2.1 Trait of leadership model . . . 119
7.2.1.1 Moving First model. . . . 119
7.2.1.2 Moving Front model. . . . 120
7.2.1.3 Reversible Agreement model. . . . 120
7.2.2 Datasets . . . 121
7.2.2.1 Simulated datasets. . . . 121
7.2.2.2 Simulation datasets for degree of hierarchy structure analysis 121
7.2.2.3 Baboon trajectories. . . . 122
7.2.2.4 Fish schools trajectories. . . . 122
7.2.3 Sensitivity analysis in model classification . . . 122
7.2.4 Hypotheses tests . . . 123
7.2.5 Parameter setting . . . 124
7.3 Results . . . 125
7.3.1 Traits of leader classification: sensitivity analysis . . . 125
7.3.2 Traits identification of baboons movement . . . 125
7.3.3 Traits identification of fish movement . . . 132
7.3.4 Traits of leaders as measure of degree of hierarchy structure . 136
7.4 Conclusions . . . 138

8 INFERRING COORDINATION MECHANISMS FROM TIME SERIES OF MOVEMENT DATA . . . 141
8.1 Introduction . . . 141
8.2 Models and properties . . . 142
8.2.1 Convergence models . . . 144
8.2.1.1 Hierarchical Model Dynamic System (HM) . . . 144
8.2.1.2 Local Reversible Agreement system (LRA) . . . 148
8.2.1.3 Discussion . . . 151
8.3 Method . . . 153
8.3.1 Setting . . . 153
8.3.2 Model fitting . . . 155
8.3.3 Model selection . . . 157
8.4 Experimental setup . . . 158
8.4.1 Simulations . . . 158
8.4.1.1 Hierarchical Model Dynamic System . . . 159
8.4.1.2 Local Reversible agreement system . . . 160
8.4.1.3 Hierarchical and Local Reversible agreement system . . . 160
8.4.1.4 Mixed strategy system . . . 160
8.4.1.5 Evaluation . . . 161
8.4.2 Baboon behavioral experiment . . . 162
8.4.3 Fish behavioral experiment . . . 162
8.4.4 Comparison with the state of the art method . . . 163

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

8.5 Results . . . 163
8.5.1 Simulations . . . 163
8.5.2 Baboon behavioral experiment . . . 166
8.5.3 Fish behavioral experiment . . . 167
8.5.4 Comparison with the state of the art method . . . 168
8.6 Conclusions . . . 169

9 MINING AND MODELING COMPLEX LEADERSHIP DYNAMICS OF MOVEMENT DATA . . . 171
9.1 Introduction . . . 171
9.2 Methods . . . 172
9.2.1 Recap of mFLICA . . . 173
9.2.2 Inferring transition diagram of leadership dynamics . . . 174
9.2.3 Mining sequence patterns of leadership dynamics . . . 176
9.2.4 Hypothesis testing . . . 177
9.2.4.1 Evaluating the significance of leadership-event order . . . 177
9.2.4.2 Evaluating the significance of frequencies of leadership-event sequences . . . 178
9.2.5 Time and space complexity . . . 178
9.3 Evaluation Datasets . . . 179
9.3.1 Type 1 Dynamics: Splitting/Merging coordination event . . . 180
9.3.2 Type 2 Dynamics: Linear coordination event . . . 180
9.4 Evaluation criteria . . . 180
9.5 Results . . . 181
9.6 Conclusion . . . 191

10 CONCLUSION . . . 192
10.1 Summary of contributions . . . 192
10.2 Assumptions and limitations of our formalization . . . 196
10.3 Future directions . . . 197

APPENDICES . . . 199
Appendix A . . . 200

CITED LITERATURE . . . 207

LIST OF TABLES

TABLE		PAGE
I	Time complexities of leadership inference methods where n is a number of time series, ω is a time window, and t^* is a length of time series.	65
II	Precision of leadership identification on simulation models. (* indicates the $std \geq 0.1$).	73
III	Initiator identification precision in fish (* indicates the $std \geq 0.1$).	74
IV	Activities classification prediction accuracy in first two days of baboon data.	78
V	Random forest classification of synthetic leadership models using proposed features	82
VI	The average difference of network density between before and after removing k individuals on the baboon dataset. We compare the case of removing high-rank individuals vs. random individuals. An element in the table represents the difference between the original network density and the network density after removing k individuals.	83
VII	The average difference of the network density between before and after removing k individuals from the fish data. We compare the case of removing high-rank individuals vs. random individuals. An element in the table represents the difference between the original network density and the network density after removing k individuals.	84
VIII	Factions and Leaders identification on simulation models	99
IX	Rank orders median accuracy within factions	100
X	A school of fish inference median accuracy over 24 trails	101
XI	Jaccard similarity between top-4 ranking individuals from centrality measures and the ground truth set of four initiators in dictatorship model from 200 datasets.	104
XII	Support of four initiators being in the list of top-4 ranking individuals from centrality measures in 100 datasets containing linear coordination events.	105
XIII	Support of four initiators being in the list of top-4 ranking individuals from centrality measures in 100 datasets containing merging/splitting coordination events.	106
XIV	Description of hypothesis tests used in this paper. A significant level has been set at $\alpha = 0.001$ for all experiments.	123
XV	PR-VCH, PR-PCH, and PR-DCH rank correlations from the baboon dataset under different thresholds.	126

LIST OF TABLES (Continued)

TABLE **PAGE**

XVI Hypothesis test results of PR-PCH and PR-DCH correlation. The zero value implies that a test fails to reject H_0 while one implies a test sucessfully rejects H_0 with $\alpha = 0.001$. 129

XVII Normal confidence intervals of PR-PCH and PR-DCH correlations from the baboon dataset at the interval level with $\alpha = 0.001$ 131

XVIII PR-VCH, PR-PCH, and PR-DCH rank correlations from the fish dataset. 134

XIX Hypothesis test results of PR-VCH, PR-PCH, and PR-DCH correlations in time-point level and interval level from the fish movement dataset. The zero value implies that a test fails to reject H_0 while one implies a test successfully rejects H_0 with $\alpha = 0.001$. 135

XX Normal confidence intervals of PR-VCH, PR-PCH. and PR-DCH correlations from the fish dataset at the interval level with $\alpha = 0.001$ 136

XXI The means and standard deviations of PR-VCH rank correlations from both simulated and biological datasets. 139

XXII The result of predicting the direction of movement via 10-fold cross validation. We compared the result of our framework (OPT) against the base-line pure strategies: HM, LRA, and AR (auto regressive strategy). (*indicates the STD $\geq 20°$) 164

XXIII The average optimal support vector $\vec{w}$ of all agents from 10-fold cross validation, inferred by our framework from simulated and the Baboon datasets. 166

XXIV Comparison between LRA and Informed strategies to predict movement directions of 24 trails of fish. 167

XXV The results of model classification of FLICA and the proposed framework via 10-fold cross validation. We use Random Forest as the main classifier for both frameworks. 168

XXVI Details of non-parametric tests used in this paper. A significant level has been set at $\alpha = 0.01$ for all experiments. 177

XXVII The example of sequences of leadership dynamics that have the highest support from HM datasets. 184

XXVIII The median of loss values in the prediction task of diagrams of leadership dynamics. 185

XXIX Hypothesis testing results of the significance of leadership-event order in Section 9.2.4.1. We reject H_0 at $\alpha = 0.01$. Each element in the table represents the percentage of the times when the tests successfully reject H_0. 186

XXX Hypothesis testing results of the significance of frequencies of leadership-event sequences in Section 9.2.4.2. We reject H_0 at $\alpha = 0.01$. Each element in the table represents the percentage of the case when the test successfully rejects H_0. 187

LIST OF TABLES (Continued)

TABLE **PAGE**

XXXI Baboons' sequences of leadership dynamics that have the top-4 highest support 189

LIST OF FIGURES

FIGURE		PAGE
1	An example of GPS-collar trajectories of Olive baboons living in Mpala Research Centre, Kenya [6, 7]. In this event, the troop is forming coordinated movement.	5
2	(Left) The example of time series U follows Q. (Right) the following network w.r.t. U follows Q relation.	24
3	U and W are sine waves that have the same frequency but different phase. U follows W with time delay $\Delta_{W \prec U}$ and W follows U with time delay $\Delta_{U \prec W}$.	26
4	(Left) the example of coordination interval in time series where $Q \prec U \prec V \prec W$. (Right) the following network w.r.t. these following relations. In this example, Q is an initiator.	27
5	(Top-left) A time series Q, (bottom-left) a time series U, and (top-right) a time series V. In this example, U partially follows Q, V partially follows U but V does not follow Q. The following network of these time series is at the bottom-right of the figure.)	29
6	Examples of individual movements from Dictatorship and Hierarchical model. Nodes represent individual positions and arrows represent directions of individual's movement. (Left) in Dictatorship model, everyone follows a leader L, while there is a hierarchy to follow for each individuals in Hierarchical model (right).	41
7	An example of following-network-density time series of Switching Dictatorship model . There are two coordination events where individual A leads a group in both pre-coordination intervals while B leads a group during coordination intervals.	42
8	An example of following-network-density time series of Event-based model. There are two coordination events where individual A leads a group in the first coordination event while B leads a group during the second coordination event.	43
9	Examples of individual movements from Initiator model. Nodes represent individual positions and arrows represent directions of individual's movement. (Left) during a pre-coordination interval, everyone follows a leader L, while, during a coordination interval, everyone knows the direction and moves to the destination directly without following a leader (right).	44

LIST OF FIGURES (Continued)

FIGURE **PAGE**

10 Examples of individual movements from Crowd model. Nodes represent individual positions and arrows represent directions of individual's movement. (Left) at time t, everyone follows some directions except informed individuals (T nodes) which moves directly to a target. Then, at time $t + l$ (right), the group's direction, which is the average of individual's directions, gradually changes toward the target. 45

11 Examples of individual movements from Linear Threshold model. Nodes represent individual positions and arrows represent directions of individual's movement. (Left) at time t, only active individuals (orange color) move toward an initiator L. Suppose $k = 3$ and $\rho = 0.50$, at time t, an inactive individual A has two active individuals F_1, F_2 and L as its neighbors. Since 66% of A's neighbors are activated, then, at time $t + 1$ (right), A is active and start moving toward L. 46

12 A high-level overview of FLICA framework 50

13 PageRank (top) and density (middle) of the 'following' network over time for an event of baboons' movement which initiates by ID3. (Bottom) The locations of individuals over three different time steps ($t = 50, 100, 250$), with the 'following' network, and PageRank indicated by node size. 51

14 (Left) Toy time series showing U following Q with a time delay $\Delta = 3$. (Right) the optimal warping path (yellow boxes) on the DTW dynamic programming matrix, shifting U backward in time onto Q. 53

15 A *coordination event* is a pair of intervals. We define the *pre-coordination interval* and *coordination interval* using threshold λ on the network density time series. 56

16 Dynamic Time Warping global vs. local example 62

17 An example of time series of following-network density. There are three coordination events where individual A leads a group in first half of all coordination events while B leads a group after A for first and second coordination events. The third coordination event has C leads the second half of the event. 67

18 (Top) NASDAQ 'following' network density and (Bottom) NASDAQ index value. Pre-coordination and coordination intervals are shown in red and green, respectively. The framework detects many known events in financial data (labeled above). Many of these events are not reflected in the NASDAQ index. 75

19 Comparison between time series of network density that generated from NASDAQ time series and from random walk time series. 76

20 Baboon activities from first two days of tracking vs. following network density 77

21 Comparison of feature spaces of leadership model classifications on simulations and real data 79

LIST OF FIGURES (Continued)

FIGURE **PAGE**

22 Loss values of prediction of the initiator support for different levels of noise and time window sizes. A lower value implies a better prediction result. 85

23 Loss values of prediction of the initiator confidence for different levels of noise and time window sizes. A lower value implies a better prediction result. 86

24 F1 scores of coordination inference for different levels of noise and time window sizes. A higher value implies a better prediction result. . 87

25 A high-level overview of mFLICA framework 91

26 Linear (above) and Splitting/Merging (below) coordination event. Each block represents a faction such that the first element is the leader ID and the second element is the member IDs set. The time interval each faction appears is at the last line. 95

27 The merging coordination event. (Top) Faction size ratios (Equation 6.1) of ID(3) and ID(18) factions. (Bottom) The GPS locations of individuals in the map over three different time steps ($t = 300, 350, 400$), with the 'following' network, and PageRank indicated by node size. ID(3) is black and ID(18) is purple. The red edges have higher edge weights than the light edges. 103

28 An example of state changing situation in the two-dimensional space. The green nodes are states of individuals at time $t-1$ and the green polygon is a neighbor-state convex hull of individual U. If U changes its state under the Bidirectional Agreement Condition, then the next state of U is always in the convex hull (orange). On the contrary, if U steps outside the convex hull (red) to make a group changes its state, then it breaks the Bidirectional Agreement Condition. 111

29 High-level overview of trait leadership scheme using FLICA ([1, Figure 1] used and modified with permission). An arrow between elements represents a relationship that an element at the rear of the arrow is the input of an element at the head of the arrow. For example, rank correlations are calculated by taking leadership ranking and convex hull ranking as inputs. 112

30 An example of position convex hull. Each point represents an individual and the polygon represents a convex hull boundary at time $t-1$. In this example, Q steps outside the convex hull at time t toward the group direction, while U steps outside the convex hull in the opposite direction. In this case, Q gets a score $+1$ and U gets a score -1 for time step t. If an individual is still in the convex hull, it gets zero score. . . . 115

LIST OF FIGURES (Continued)

FIGURE		PAGE
31	Sensitivity analysis in model classification task from simulation datasets with different noise levels. (Top) 10-fold cross validation loss values. Each element in the table represents the loss value of each each noise setting (γ, β). (Buttom) Rank Correlation between actual leadership ranking vs. predicted ranking from moving first and moving front models.	127
32	Comparison of PR-VCH, PR-PCH, and PR-DCH rank correlations under different thresholds. For PR-VCH correlation, the results in both time-points level (Top-left) and interval level (Bottom-left) show that there are no strong correlations between leadership and VCH ranking. In contrast, leadership and PCH rankings have positive correlations in both time-points level (Top-middle) and interval level (bottom-middle), as well as PR-DCH correlation has a negative correlation at the interval level (bottom-right).	128
33	Comparison of PR-VCH, PR-PCH, and PR-DCH rank correlations in both time-point and interval levels from the fish movement dataset. In the time-point level (left), the result shows that leadership vs. VCH, and leadership vs. PCH rankings have positive correlations, while leadership and DCH has negative correlation. In the interval level (right), leadership vs. VCH and leadership vs. PCH rankings have stronger positive correlations than time-point level, while leadership and DCH also has stronger negative correlation.	133
34	The distributions of PR-VCH rank correlations of datasets from three leadership models. The higher rank correlation implies the higher degree of hierarchy structure of early movement order in the model.	137
35	An example of communication networks $G(t) = (\mathcal{V}, E(t))$ between $t = 0$ and $t = 3$ (above). These networks are the realization of the probabilistic following network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ (below). The arrows represent the directed edges while the dashed lines are empty edges. When the time step increases, the informed agent L can increasingly spread its state (orange node) to more follower nodes (blue nodes).	145
36	An example of physical points as positions and state points as directions. In position space (above), the individual i (red node) has all gray nodes as its neighbors in LRA since they are neighbors in Delaunay triangulation. In the direction space (below), i updates its next direction to be A rather than B since B is outside the i's neighbor convex hull. .	149

LIST OF FIGURES (Continued)

FIGURE **PAGE**

37 An overview of the proposed framework. Given a set of time series as inputs, first, the framework detects coordination intervals. Second, the framework infers the optimal strategy from a set of candidate strategies. In the model fitting step, the framework evaluates the performance of each pure and mixed strategy on the task of individual-level direction prediction using training datasets, then it gives the support weight to each strategy w.r.t. some minimum weight threshold $\vec{\kappa}$. The threshold $\vec{\kappa}$ represents the weight bias toward specific strategies. In the model selection step, the framework selects the best strategy that performed well in validation datasets from a set of various strategies that have been trained by different $\vec{\kappa}$. Finally, the framework reports the optimal support weight of each strategy for each individual. If the weight w_i of strategy f_i is high, then i might use f_i strategy to coordinate with the group. For simplicity of the exposition, in this paper, we deploy three candidate strategies: 1) an individual moves following specific individuals (HM), 2) an individual moves following its physical neighbors (LRA), and 3) an individual moves toward the same direction as in its previous time steps (Auto Regressive). However, other candidate strategies are also admissible. 152

38 An example of movement strategy inference for i. Given the information on positions and directions of individuals in the past (blue and green nodes), we want to infer the i's strategy of movement that can be whether that i's next direction follows its neighbors (A node), or follows specific individuals (B node), or neither (C node). 153

39 A high-level overview of the proposed framework. 173

40 Splitting/Merging (above) and Linear (below) coordination event. Each node represents the ID of leader of each sub-group at the particular time and each edge represents the change of group's leaders. 179

41 The example of the inferred diagram of leadership dynamics by our framework from a Type-1-HM dataset. Comparing the inferred diagram with the ground truth, only nodes $\{2, 4\}$ and $\{2, 3\}$ are false positive nodes. The support of $\{1\}, \{2, 3, 4\}, \{3\}$ and $\{4\}$ should be 0.25, and our framework can infer the support for each node closely to 0.25. 182

42 The example of the inferred diagram of leadership dynamics by our framework from a Type-2-HM dataset. Comparing the inferred diagram with the ground truth, there are no false positive nodes. The support of $\{1\}, \{2\}, \{3\}$ and $\{4\}$ should be 0.25, and our framework can infer the support for each node closely to 0.25 183

FIGURE **PAGE**

43 The inferred diagram of leadership dynamics of the baboon dataset from our framework. Each row represents the node of leader sets of previous state and each column represents the next state. Each row label consists of baboon gender: 'M' or 'F', a set of frequent-leader IDs, and the support value of frequent-leader set. For example, in row 3 and column 2, the event that two female baboons F18 and F22 are leading their separate sub-groups concurrently can happen with the support 0.1 (out of all the coordination times). These two sub-groups have a chance to be merged together to a larger group lead by F18 with the probability 0.29. 190

LIST OF ABBREVIATIONS

GPS	Global positioning system.
IM	Influence Maximization (model).
DM	Dictatorship Model (model).
HM	Hierarchical Model (model).
LRA	Local Reversible agreement system (model).
LT	Linear Threshold Model (model).
IC	Independent Cascade Model (model).
CM	Crowd Model (model).
FLICA	Framework for Leader Identification in Coordinated Activity (framework).
mFLICA	Framework of Multiple-Faction Leadership Inference in Coordinated Activity (framework).
HMM	Hidden Markov Model (model).
DTW	Dynamic Time Warping (similarity measure).
VCH	Velocity Convex Hull.
PCH	Position Convex Hull.
DCH	Direction Convex Hull.
PR	PageRank (graph centrality measure).

SUMMARY

When a group of people decides to move somewhere together, who is the initiator who starts moving and everyone follows? Do the group members follow friends around them or do they prefer to follow specific individuals? These questions are about leadership. Leadership plays a key role in social animals', including humans', decision-making and coalescence in coordinated activities such as hunting, migration, sport, diplomatic negotiation, etc. In these coordinated activities, leadership is a process which organizes interactions among members to make a group achieve collective goals. Understanding initiation of coordinated activities allows scientists to gain more insight into social species' behaviors. However, by using only the data on time series of activities, inferring leadership, as manifested by the initiation of coordinated activities, faces many challenging issues. First, there is no fundamental concept to describe these activities computationally. Second, coordinated activities are dynamic. Third, several different coordinated activities may occur simultaneously among subgroups. To fill these remaining gaps in leadership inference, we formalize several computational leadership problems and propose methodologies to solve them.

First, we formalize the Coordination Initiator Inference Problem and proposed a simple yet powerful framework, FLICA, for extracting periods of coordinated activity and determining individuals who initiated this coordination, based solely on the activity of individuals within a group during those periods. Second, we extend a concept of coordinated activities to allow them to overlap as multiple coordinated factions. We formalize Faction Initiator

Inference Problem and proposed a leadership inference framework as a solution of this problem. The framework makes no assumptions about the characteristics of a leader or the parameters. Third, we present a computational method to characterize and classify the traits of leaders in movement initiation. We propose a framework for ranking leaders according to their position, velocity, and heading relative to the group and perform hypothesis testing of correlations between target features and leadership ranking. Fourth, we propose a methodology to infer mechanism of coordinated activity. Given a time series that include coordinated movement and a set of candidate strategies as inputs, we provide the methodology to infer the set of strategies that each individual uses to achieve movement coordination at the group level. Lastly, we focus on mining and modeling frequent patterns of leadership dynamics. We formalize a new computational problem, Mining Patterns of Leadership Dynamics, as well as propose a framework as a solution of this problem. Our framework can be used to address several questions regarding leadership dynamics of group movement.

We evaluate and demonstrate the performance of the proposed frameworks in both simulated and real-world datasets, such as baboon trajectories, time series of fish movement as well as time series of closing price of stock market. The frameworks perform better than non-trivial baselines in both simulated and real-world datasets. Our problem formalization and frameworks enable opportunities for scientists to analyze coordinated activities and generate scientific hypotheses about collective behaviors that can be tested statistically and in the field.

CHAPTER 1

INTRODUCTION

Which zebra initiated the flight from a lion? Whom does the elephant herd follow to water? Who is the trend-setter whose opinion many follow at the moment? (And is it the same person whether it's the opinion about the future of AI or the hottest lunch spot?) In all these scenarios, the leader might not be the one who is speaking the loudest or positioned at the front of the group *after* the group has already agreed to follow [8, 9]. Thus, in order to identify those leaders or trend-setters, we must also determine the moment of the group's decision to follow as well as the dynamics of group coordination. In this work, we present theoretical properties, methodologies, and methods to analyze leadership of coordination based on time series data. Our work generalizes the concept of leadership as an initiation process, which can be applied to any domain including ecology, social science, economics, AI, etc. We demonstrate the applications of our proposed frameworks in simulation, animal datasets, as well as a stock market closing price dataset. In addition, our problem formalization and frameworks enable opportunities for scientists to analyze coordinated activities and generate scientific hypotheses about collective behaviors that can be tested statistically and in the field.

1.1 Motivation

When problems cannot be solved by a single individual, group members have to form coordination. One of methods that makes a group to coordinate with each other is through

'leadership'. Leadership is a process of specific individuals (leaders) affecting group's actions and decisions in order to make a group coordinate to achieve collective goals [10, 11].

Leadership plays a key role in solving collective-action problems (e.g. social conflicts, migration, hunting, territorial defense) across social species [11], organizing the collective (i.e. group) behaviors of social organisms ranging from humans [8] to hymenoptera [12]. It potentiates complex patterns of cooperation and conflict (e.g., lions [13], hyaenas [14], meerkat [15], chimpanzees [16], humans [17]), organizes group movements (fish [18], humans [8]), and may prevent free-riding [19]. In the context of coordination, which is defined as an emergence of collective actions to achieve the collective goals [20], leadership mainly contributes by fostering collective behaviors in social species ranging from humans [8, 10, 11] to fish [21].

The availability of data from physical proximity sensors, GPS, and the web open up the possibility of measuring leadership as the process of initiation in online activities, face-to-face human interactions, animal populations, and aggregate social processes such as economic activities. However, these new types of data have unique characteristics and require new types of computational techniques to analyze them.

1.2 Related work

Coordinating patterns of individual activity is a challenge that all social organisms face. In the context of group behavior and decision-making in biology and sociology, leaders are individuals who successfully induce a group of others to follow them to a common goal, state, or behavior [18, 22–24]. Biological studies showed that leaders may be context-specific [18, 25]

and the important initiators of particular group activities are not necessarily the individuals found at the top of their group's social dominance hierarchy [7, 26].

Substantial interest currently exists in identifying leaders and determining how they influence the behavior of others in their social environment. Previous work in several domains defined leadership according to physical characteristics (*e.g.,* size, sex [22]), positions in physical movement in public spaces [27], location-based social networks [28], rule-based models [25], physical trajectory, and association patterns [29, 30].

Computationally, the majority of previous works use a global notion of leadership and creates a global, static leadership ranking over the entirety of the input data [31, 32]. Other domain-specific methods infer leadership from implicit pairwise dyadic dominance or leader-follower interactions [28, 30, 33]. Some methods define an explicit network over the dyadic interactions or use a known network topology [34] and use network measures, such as PageRank and HITS, or cascade size to identify leaders [32]. In this work, we focus on leadership in the aspect of computation.

1.2.1 Computational leadership categories

Among the computational approaches, leadership frameworks can be separated into two main types: movement leadership and Influence Maximization.

1.2.1.1 Movement leadership

The early pioneer work on trajectories leadership was done by Andersson *et al.* [30], which defines leadership based on the flock patterns in [35]. Leaders have to satisfy three conditions. First, a leader has to have enough followers. Second, a leader must continuously lead its

followers long enough. Third, a leader never *follows* others during the leading time. Therefore, there is only a sole leader for each specific period of time. According to Kjargaard *et al.* [33], this framework can easily fail in the case of noise. Moreover, it cannot detect a leader when a group of people walks through a crowd, since there might be other people who are in front of the group's leader. Hence, it is not a realistic framework, since a leadership model should consider not just only positions of individuals.

The frameworks by Kjargaard *et al.* [33] and Solera *et al.* [27] are developed based on the work by Andersson *et al.* [30]. The difference between them is that [27, 33] use a following network to analyze leadership instead of using a flock pattern. The benefit of using a following network is that the frameworks in [27, 33] can provide a leadership ranking of individuals and can consider the group hierarchy of following instead of reporting only who is a leader, as in [30].

1.2.1.2 Influence maximization

Leadership has also been studied in explicit social network settings. From a social network perspective, leaders can be characterized as influential individuals who have many followers that imitate the leader's actions [31], and, thus, successfully take a group from one behavioral state to another. Much of the computational work has focused on the problem of influence maximization (IM)–i.e. how individuals are able to maximize their impact on the behavior of the group as a whole by iteratively affecting local network neighborhoods [36,37]. This approach assumes that the network structure is known.

The work by Goyal *et al.* (2008) [31] deals with the social network in the same way as [37]. However, instead of solving the Influence Maximization problem, the framework by Goyal *et*

al. (2008) uses influence relationship to determine a set of leaders. Even though the frameworks from [31] and [30] deal with different data, they share the same concept of leadership; a leader [30] or genuine leader [31] is an individual who has sufficiently many followers and never *follows* others.

1.2.2 Coordination mechanisms inference from time series

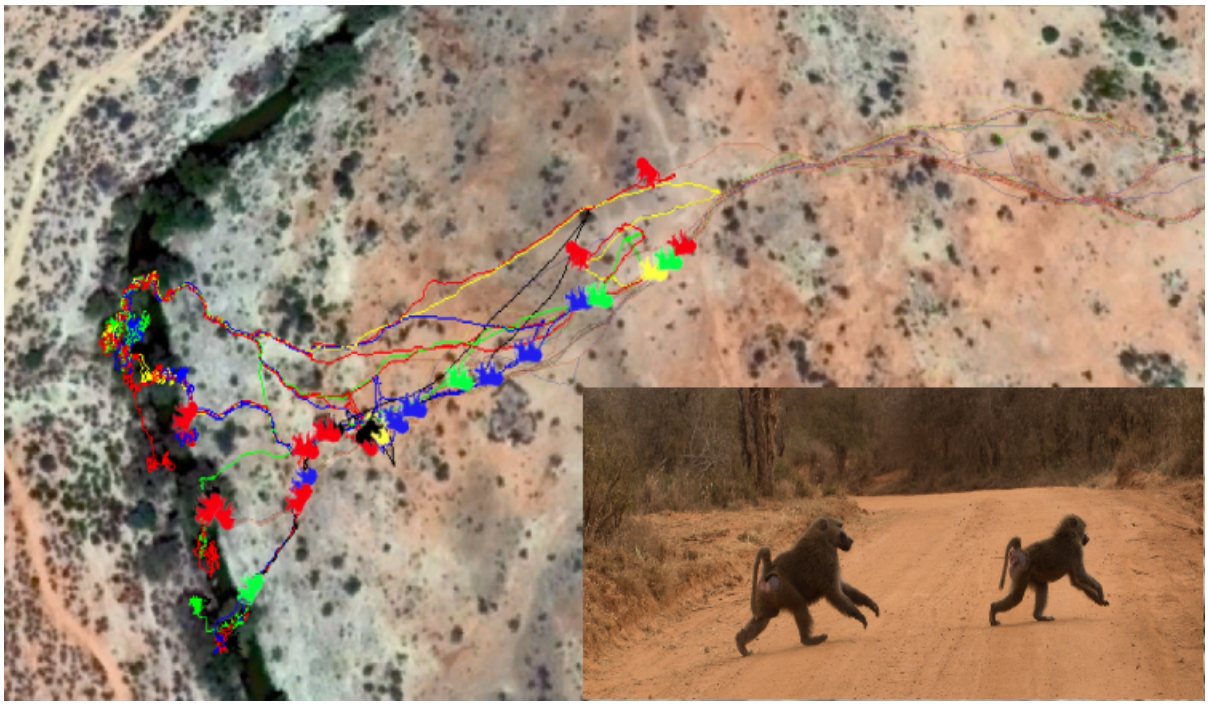


Figure 1. An example of GPS-collar trajectories of Olive baboons living in Mpala Research Centre, Kenya [6, 7]. In this event, the troop is forming coordinated movement.

Coordination is a form of group behavior aimed to make the group achieve a collective goal [20,38,39]. During the decision making process, a collective goal is to reach a group's consensus, which is defined as the state when all individuals share a common agreement [40]. One of the mechanisms by which a group can achieve a collective goal is leadership, which is a process of pattern initiation by specific individuals, leaders, then followed by the rest [1]. In behavioral studies, coordination problems, such as group decision making, coordinated movement, group hunting, social conflicts, and territorial defense, can be solved by leadership [8, 11, 21, 41–43]. However, leaders might not be explicit or global to a group, yet the group can still create coordinated movement via a local strategy (e.g. individuals follow their neighbors) [8, 41, 44].

In cooperative control of multi-agent systems, the field focuses on how to design a local strategy for each agent so that the group can achieve collective goals [40, 45]. Agents can communicate only with their neighbors via a communication network, which is defined by any neighborhood concept in some space [45]. There is a large body of work in multi-agent systems that proposes local synchronization strategies [40, 45]. In behavioral studies, the work in [8,46] tries to model the coordination process via a concept of information spreading. A small number of informed agents can spread information through a large number of uninformed agents, which results in the group's consensus and coordinated movement. The work by Chazelle [47] introduces a model, namely a reversible agreement system, that guarantees convergence of the group state, with or without leaders. In online social networks, there is also the "Diffusion Model" [36,37,48] that models an information spreading process among individuals that results in the entire network reaching a common state.

However, in this work, we focus on the inverse question of *inferring* the local strategies collective individuals use to achieve a state of coordination. There are only a few studies that address this question. The works by Farine *et al.* [49, 50] found that wild baboons can achieve the state of coordinated movement within a group by following their neighbors or long-term associates, depending on the time scale of the coordination process. There are several studies that look at the collective behavior of fish. For example, the work in [51] models and infers the rules of movement coordination of fish, which is affected by the group size; Herbert-Read *et al.* [52] report that the rules of movement coordination of fish mainly depend on attraction forces of the group; and Katz *et al.* [53] show that fish tend to imitate the direction of neighbors ahead. The work in [54, 55] proposed model selection methods to infer the behavior model of animals, but the method cannot be used to find models that guarantee coordination. Note that we use the words 'model', 'mechanism', and 'strategy' interchangeably.

1.2.3 Mining and modeling frequent patterns of leadership dynamics

Leadership is an essential part of collective decision and organization in social animals, including humans. In nature, leadership is dynamic and varies with context or temporal factors. Understanding dynamics of leadership, such as how leaders change, emerge, or converge, allows scientists to gain more insight into group decision-making and collective behavior in general. However, given only data of individual activities, it is challenging to infer these dynamic leadership events.

For example, suppose i and j lead separate sub-groups, how often do the two groups merge to a larger group lead by k? How likely is it that the group lead by i will split into more than three sub-groups?

There are many works focusing on inferring dynamics of groups or clusters [56,57]. However, mining frequent patterns of leadership dynamics requires an approach to identify both groups and leaders of those groups that change over time. Moreover, since the groups following a leader during coordination have a special structure of following relations among members, the standard clustering methods cannot be used in this case.

1.2.4 Limitations of current methods

In this work, we consider leaders as initiators who initiate patterns of action and everyone within a group follows. In reality, first, there are multiple coordinated activities happening within a single set of time series. Second, for each coordination, there might be multiple sub-coordination events as well as multiple initiators appearing simultaneously. Given only time series and no further information, the problem of leadership inference is far from trivial. Below is the list of issues that existing methods cannot solve.

- **Coordination Model:** In movement leadership, existing works mainly focus on identifying leaders with some set of rules or assumptions, but there is no work considering the leadership model inference based on time series.

  In Influence Maximization, the majority of papers focus on Linear Threshold and Independent Cascade models as main coordination mechanisms. However, there are other models that can

be represented as a coordination mechanism, such as Dictatorship and Hierarchy models, as well as non-network based models. Given time series of group activities as an input, there is no existing methodology to infer leadership mechanism behind group's activity.

- **Coordinated Activity:** In movement leadership, flock patterns are used to detect group movement. However, since a flock pattern is defined based on a simple set of rules, it cannot be used to detect dynamics of coordinated activity that break its simple rules. For example, a group of individuals might have coordinated movement with arbitrary shape and radius, but a flock pattern can detect only group movement pattern that have a fixed radius with an eclipse-like shape. Generally, coordinated movement can be non-linear in directions and positions of movement. Since a flock-based model has assumptions about a shape of group movement and a positions of leader, who must always be at the front of the group, therefore the model cannot detect any complicated coordinated movement in general.

Influence Maximization focuses on only one kind of state changing, which is an information spreading event. In a coordinated activity, any state that the group coalesces to, a posteriori, is a coordinated state, and it may change from one time to another. Coordinated activity can represent multiple types of coalesced states over multiple time points. For example, a coordinated movement of animals to a feeding site is considered to be a coordinated activity for that group, but so is sudden jumping up and down in agitation, or all looking in the same direction, or all falling sick. And the coordinated movement can be different and look different

in the morning versus in the evening. There is also no computational method that can be used to infer coordinated activities and their initiators from time series of group activities.

- **Coordination Based on Context:** In movement context, a flock pattern is used to identify coordinated movement pattern with a set of simple conditions, however, by using only time series and no future information, a flock-like model also fails to detect multiple coordinated movement that have different initiators who are not always in the front. Moreover, coordinated movements that exist in time series may not satisfy any simple conditions of a flock pattern, hence a flock-like framework might be unable to detect any coordinated movement at all.

  In social network context, the Influence Maximization assumes that there is only one coordination event of information spreading. However, in reality, information spreading or coordination events can occur many times. For instance, within a day, a group of animals can have many coordinated movement events to many places.

In summary, there is a clear gap between the biosociological definitions of leadership in group decision-making and the existing computational approaches. Currently, there are no computational approaches that (1) view leaders as initiators of group behavior change, which can (2) identify the timing of the process of the change initiation and the group's decision-making in (3) arbitrary contexts, under (4) a variety of leadership models, (5) identifying traits of leaders, as well as (6) mining and modeling frequent patterns of leadership dynamics.

1.3 Our proposed solutions

In the first part of this work, we present the new computational problem of inferring leader identity in the context of successful initiation of coordinated activities among groups of individuals or other entities (Section 2.1 and 3.1), as well as proposes the first automated method for unsupervised leader identification (Chapter 5). The method uses only time series activity data of entities, with no additional information. The proposed approach automatically determines (1) the time interval of group coordination, (2) the time when the (possibly implicit) decision for that coordinated activity was made, (3) the identity of the coordination initiator, and (4) the mechanism by which the group came to follow the initiator by classification methodology.

In the second part of this work, we extend the concept of coordinated activity from the first part to *multiple* coordinated activities which can occur simultaneously (Section 2.2 and 3.3) as well as propose the framework to analyze these complicated coordinated activities (Chapter 6). The new proposed method can determine (1) the time interval of multiple subgroup coordination events, (2) the identity of these coordination initiators, and (3) the dynamics of coordination events (group merging or splitting).

In the third part of this work (Chapter 7), we propose the concepts of traits of leaders in time series as well as provided a framework to infer traits of leaders of coordinated movement.

In the fourth part of this work, (Section 2.4, Section 3.5, and Chapter 8) we propose the new concept of inference of coordination mechanism that considers a convergence property of models and other properties since the mechanism inference by the existing classification methodology has many limitations to capture the characteristics of leadership model. For example, by

representing leadership mechanisms as features for classification, we cannot consider whether leadership models possess a convergence property that creates coordination. The proposed framework is able to infer the set of strategies that each individual uses to achieve coordination at the group level.

Lastly, in Chapter 9, we proposed a framework that is capable of mining and modeling frequent patterns of leadership dynamics from multiple subgroup coordination events.

CHAPTER 2

PRELIMINARIES

In this dissertation, it consists of the works from five previous works: the work in [1–3] (Section 2.1 and Chapter 5), [4] (Section 2.2 and Chapter 6), [5] (Section 2.3 and Chapter 7), Chapter 8 (Section 2.4), and [58] (Section 2.5 and Chapter 9). First, in FLICA framework [1–3], we established fundamental concepts regarding leadership of coordination and decision-making intervals inference as well as leadership model selection classification method. Second, in mFLICA framework [4], we extended the concept of leadership from a single group of coordination in FLICA to the concept of multiple factions in [4] that allows the sub-coordinated activities happen simultaneously. We also proposed time window inference in [4] based on the concept of coordination measure. Third, in [5], we proposed the concepts of traits of leaders in time series as well as provided a framework to infer traits of leaders. Fourth, we proposed the coordination mechanism inference concept and a framework to infer an underlying model of collective behaviors of each individual. Finally, we proposed a framework for mining and modeling frequent patterns of leadership dynamics from multiple subgroup coordination events inferred by the mFLICA framework [4].

2.1 Coordination event detection and initiator identification in time series

COORDINATION INITIATOR INFERENCE PROBLEM: An agreement of a group to follow a common purpose is manifested by its coalescence into a *coordinated* behavior. The process of initiating this behavior and the period of decision-making by the group members necessarily precedes the coordinated behavior. **Given time series of group members' behavior, the goal is to find these periods of decision-making and identify the initiating individual, if one exists.**

The first part of my contribution, which comes from the work in [1, 2], is establishing and formalizing the **new computational problem of coordination initiation inference**. We call it the COORDINATION INITIATOR INFERENCE PROBLEM. The formulation is a **generalization** of many related leadership and initiation inference computational problems. We explicitly relate existing leadership and influence propagation problems as special cases of our formulation. The new formulation uses only the time series of individual behavior as input, with no assumption of additional information such as demography, prior history, dominance hierarchy, or a network structure. The problem formulation aims to identify different local instances of behavior initiation, allows the identity of the initiator to be instance-specific, and makes no assumption on the leadership or behavioral model.

Our additional contribution is in proposing a computational **solution framework** to this new COORDINATION INITIATOR INFERENCE PROBLEM. We propose a general, scientifically grounded, unsupervised, and extendable framework with few assumptions for identifying indi-

viduals who lead a group to a state of coordinated activity (or, more generally, an entity that induces group coalescence). Our framework is capable of:

- **Detecting coordinated activity events:** discovering coordination intervals and decision-making periods leading to that coordination;
- **Identifying initiators:** identifying the initiators of this coordinated behavior, that is, the individuals who succeeded in leading the group to coordination, specifically locally to each coordination instance; and
- **Classifying the group coordination model:** characterizing the type of the group's transition behavior to coordination according to interpretable, dynamic models.

We demonstrate the framework's ability to analyze leadership in coordinated activity on synthetic and real datasets over several domains. We compare our framework with state-of-the-art methods for leadership identification for the special cases of our problem where such methods are applicable. For many instances of our new problem, there are no existing methods. We demonstrate that existing solutions fail and do not extend to these instances. We use synthetic simulated data to validate each aspect of the framework. We analyze two biological datasets – GPS tracks of a baboon troop and video-tracking of fish schools, – as well as stock market closing price data of the NASDAQ index. The results are consistent with ground-truthed biological data. Moreover, the framework finds many known events in financial data, which are not otherwise reflected in the aggregate NASDAQ index. Our approach is easily generalizable to any coordinated activity in time series data of interacting entities.

2.1.1 Influence maximization vs. coordination initiator inference problem

The Influence Maximization problem is closely related to Coordination Initiator Inference Problem. In fact, *successful* Influence Maximization, where a large fraction of the population is influenced, is a special case of Coordination Initiator Inference Problem. When the influence is spread to a majority of the population, that population is now in a coordinated state, with the decision period starting at the initiation of the influence and the initiators being that source of influence. However, Coordination Initiator Inference Problem goes beyond Influence Maximization in every aspect of the framework and can capture different models of decision-making, coordinated activity, as well as repeated and context-specific dynamics of coordination.

- **Coordination Model:** Our new problem formulation, Coordination Initiator Inference Problem, generalizes to all types of models that can be represented as a coordination mechanism. To recognize difference types of coordination mechanisms, we also provide the model classification approach to classify these coordination models based on some proposed features.
- **Coordinated Activity:** In Coordination Initiator Inference Problem, the problem focuses on not only which individuals initiate coordinated activities, but also **when** coordinated activities occur, without the explicit prescription of the type of coordinated activity. To analyze coordinated activities in time series, we provide a framework to detect coordination intervals as well as initiators of these coordination events.

- **Coordination Based on Context:** The original Influence Maximization assumes that there is only one coordination event of information spreading. However, in reality, information spreading or coordination events can occur many times. For instance, within a day, a group of animals can have many coordinated movement events to many places. Our work here aims to infer the multiple coordination events, which might have different initiators.

2.2 Inferring leadership dynamics of complex movement from time series

> FACTION INITIATOR INFERENCE PROBLEM: To reach collective goals, group's members must coordinate with each other. Multiple factions within a big group may exist solving their sub-tasks in helping the entire group achieve the collective goals. **Given time series of individual activities, our goal is to identify periods of coordination and the subsequent coordinated activity, find factions of coordination if more than one exist, as well as identify leaders of each faction**

In the context of coordination leadership, the approach of leadership inference in [1] provides a solution for identifying coordination events, the initiators of these events, as well as proposes an approach for the classification of the types of leadership models acting on a group. However, the framework in [1] cannot be used to infer *multiple* coordinated activities which can occur simultaneously because the notion of multiple factions is not employed by the framework. Our main purpose is to close this gap in the study of leadership of coordination.

First, we introduce the novel computational problem of leadership identification in multiple coordinated activities, namely FACTION INITIATOR INFERENCE PROBLEM. We formalize the

problem and analyze its theoretical properties and implications. Second, we propose a high-performance framework for Faction Initiator Inference Problem by combining several existing methods in a principled and novel manner. Our framework is capable of:

- **Detecting intervals of multiple coordination:** inferring intervals when different coordinated activity in multiple groups may appear simultaneously;
- **Identifying leaders:** identifying the initiators of these coordinated activities, the individual who initiates each coordination and the group that follows; and
- **Discovering the events of merging and splitting of coordination:** identifying the time when a coordinated group is separated into smaller sub-groups or merged with another coordinated group.

We demonstrate the ability of the framework to infer leadership in multiple coordinated groups on both simulated and biological datasets. Since we propose the new problem and framework and no other approaches exist, we compare our framework against a non-trivial baseline, which is the modification of the closest existing approach in leadership inference. Our approach is flexibly generalizable to any multiple coordinated activities from any time series data.

2.3 Identifying traits of leaders in movement initiation

In this work in [5], we propose an explicit framework for testing hypotheses about the behavioral traits of leaders by combining leader identification approaches with leadership characterization. We first identify leaders and then evaluate behavioral traits that purportedly characterize a leader. In the context of movement initiation, we focus on three behavioral traits

commonly assumed to be associated with leadership of group movements: 1) being at the front of the group [30], 2) being the first to start moving [41], and 3) being the first to move in the new direction [59]. The framework is general enough to incorporate any set of traits as the set of hypotheses of leader characterization.

LEADER TRAITS CHARACTERIZATION PROBLEM: **Given a time series of individual activities and target traits, the goal is to find a set of leaders during decision-making periods and a set of traits that best characterize these leaders.**

We propose a two step approach for the LEADER TRAITS CHARACTERIZATION PROBLEM:

1. We find instances of leadership and identify leaders, using an agnostic and assumption free leadership inference framework FLICA [1];
2. We evaluate traits of interest for all identified leaders and perform hypothesis testing to infer which traits are significant.

We demonstrate our approach using a publicly available position tracking data set from a troop of wild olive baboons (*Papio anubis*) from Mpala Research Centre, Kenya [6, 7]. Our results show that, in baboons, movement initiators are not the first to move but, instead, are the first to explore new areas, and that the group quickly aligns itself with the direction of the leader's movement.

2.4 Coordination mechanisms inference from time series

COORDINATION STRATEGY INFERENCE PROBLEM: To reach a group consensus, individuals have to coordinate with others. There are many strategies each individual can use to achieve coordination at the group level. **Given time series of individual activities and a set of candidate strategies, the goal is to find the set of original strategies individuals used that lead to the group consensus.**

The work in [1–3] (Section 2.1) provided a framework, FLICA, for leadership inference and model classification in time series data. FLICA considers the shape of time series to infer pairwise relationship who follows whom (instead of considering only directions or positions of individuals). Hence, FLICA subsumes all previous methods [3] including FLOCK patterns leadership [30, 60], time-lag following leadership [33], etc. FLICA can infer an underlying possible group model that generated coordination via a classification method. However, FLICA cannot be used to infer individual-level strategies that collectively combine to coordinated movement at the group level. In fact, each individual within a group can use a *different* strategy to achieve collective coordination (Proposition 8.2.2).

In order to fill the gap in the literature, we formalize COORDINATION STRATEGY INFERENCE PROBLEM, analyze theoretical properties of a strategy that guarantees coordination, propose hierarchical and non-hierarchical strategies that guarantee coordination, as well as propose a computational framework to infer, from time-series data, individual-level coordination strategies.

Given a set of candidate strategies and time series of coordinated movement, our framework is capable of:

- **Inferring the latent strategies:** inferring the best fit set of mixed or pure strategy for agents that provide the lowest loss value for the task of predicting the direction of movement; and
- **Movement prediction:** predicting the direction of the next move of each agent when the optimal strategy is unknown, using the set of the inferred latent strategies.

We evaluate and demonstrate the performance of our framework on simulated datasets as well as real-world datasets of animal movement. On simulated data, the task is to infer the correct latent coordination model that was used to generate the simulated time series of coordinated movement. We use the baboon dataset to predict the next movement to find which strategies each baboon likely used to coordinate its movement. Lastly, in fish datasets, we show how to apply the framework to do the model selection to address a hypothesis about the original model that the fish use to achieve coordinated movement.

2.5 Mining and modeling complex leadership dynamics of movement data

In this work [58], we formalize the problem of Mining Patterns of Leadership Dynamics, as well as propose a framework, which is the extension of mFLICA [4], as a solution to this problem. We adapt the traditional framework of frequent pattern mining [61–63] and the Hidden Markov Model (HMM) approach [64] to model the dynamics of frequent patterns of leadership. Our framework is capable of:

MINING PATTERNS OF LEADERSHIP DYNAMICS: Given time series of individual activities, the goal is to mine and model frequent patterns of leadership dynamics, including emergence of multiple leaders, convergence of multiple leaders to a single one, or change of a leader.

- **Mining and modeling frequent patterns of leadership dynamics :** inferring the transition diagram of frequent dynamics of complex leadership events, such as "a single group lead by k splits into two groups lead by i and j". In addition, we infer the probabilities of the transitions between such two events in the diagram.
- **Evaluating the significance of leadership-event order:** we propose a null model of the dynamics of leadership events and perform hypothesis testing to compare frequent-pattern model of leadership dynamics inferred from the given input to our proposed null model.
- **Mining sequence patterns of leadership dynamics:** finding support values for the leadership-dynamics sequences from time series of movement data.
- **Evaluating the significance of frequencies of leadership event sequences:** we propose a null model of the sequences of leadership events and perform hypothesis testing to compare the support distribution of leadership event sequences inferred from the given input to our proposed null model.

We evaluate our framework performance by using several simulated datasets, as well as using the real-world dataset of baboon movement to demonstrate the application of our framework. There are no existing methods to address this problem, thus, we modify and extend

the existing leadership inference framework to provide a non-trivial baseline. Our framework performs better than this baseline in all datasets. Moreover, we also propose a method to perform statistical significance tests, comparing inferred frequent patterns of leadership dynamics with our proposed null hypotheses. Our framework opens the opportunities for scientists to generate scientific hypotheses that can be tested statistically regarding dynamics of leadership in movement data.

CHAPTER 3

PROBLEM FORMALIZATION AND ANALYSIS

3.1 Single coordination event without noise

Given a collection of time series, we want to find initiators of highly coordinated patterns. To formally state the COORDINATION INITIATOR INFERENCE PROBLEM, we need to formalize notions of "coordination" and "initiation."

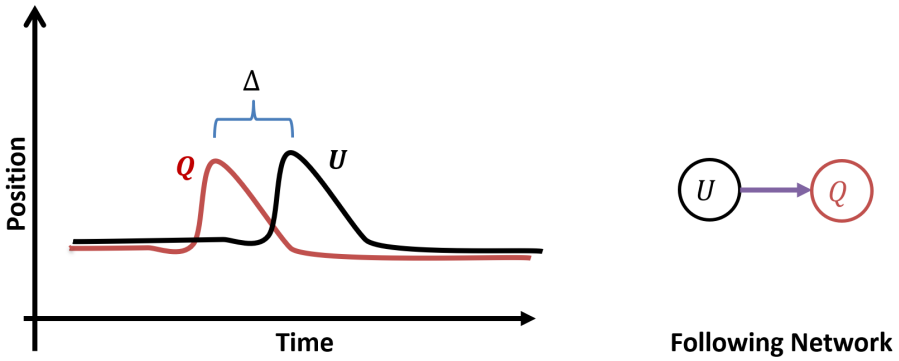


Figure 2. (Left) The example of time series U follows Q. (Right) the following network w.r.t. U follows Q relation.

First, we define an intuitive notion of a FOLLOWING RELATION, as "two individuals performing the same sequence of actions (or generating time series values) with some fixed delay

(Figure 2)." In this work, given a m-dimensional time series U, we use $U(t)$ to refer to an element of the time series U at time t and, for a given $\Delta \in \mathbb{Z}^+ \cup \{0\}$, U_Δ as a time-shifted version of U where, $U(t) = U_\Delta(t + \Delta)$. Formally:

Definition 1 (FOLLOWING RELATION) *Let U and W be m-dimensional, arbitrary-length time series. If for all $t \in \mathbb{N}$, there is a fixed time delay $\Delta \in \mathbb{Z}^+ \cup \{0\}$ such that $W(t) = U(t + \Delta)$, then U follows W denoted as $W \preceq U$. We denote $W \prec U$ if $\Delta > 0$.*

Lemma 3.1.1 *Let U and W be time series such that $W \preceq U$ and $U \preceq W$, then U and W are equivalent time series denoted $U \equiv W$.*

Proof There are two cases when both $W \preceq U$ and $U \preceq W$. First, $W = U$ and $U = W$ (that is, $\Delta = 0$ in both following relations). Clearly, $W \equiv U$. Second, $W \prec U$ with $\Delta_w > 0$ and $U \prec W$ with $\Delta_u > 0$. Then, by definition, $W(t) = U(t + \Delta_w)$ and $U(t) = W(t + \Delta_u)$. Therefore, $W(t) = W(t + \Delta_w + \Delta_u)$. Thus, if $W \preceq U$ and $U \preceq W$ then $W \preceq W$ (and similarly $U \preceq U$). Thus, the two time series are identical periodic with a different starting point and therefore equivalent.

In periodic time series, such as a sine wave, we use Lemma 3.1.1. If we have two sine waves that have the same frequency but different phase, then we consider them as a pair of time series that follow each other, from which it can be concluded that they are equivalent. For instance, in Figure 3, U and W are two sine waves, which have the same frequency but different phase. Because $U \prec W$ with time delay $\Delta_{U \prec W} > 0$ and $W \prec U$ with $\Delta_{W \prec U} > 0$, by the second case in Lemma 3.1.1 above, $U \equiv W$.

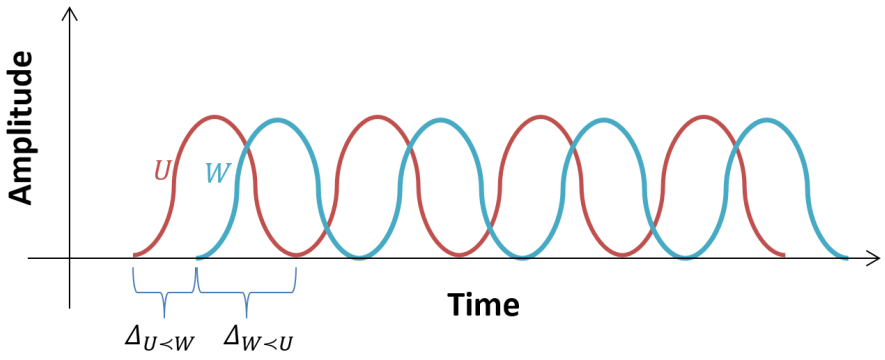


Figure 3. U and W are sine waves that have the same frequency but different phase. U follows W with time delay $\Delta_{W \prec U}$ and W follows U with time delay $\Delta_{U \prec W}$.

Lemma 3.1.2 *The* FOLLOWING RELATION *is a partial order over time series [65].*

Proof Antisymmetry: if $W \preceq U$ and $U \preceq W$, then $W \equiv U$ by Lemma 3.1.1. The FOLLOWING RELATION is also trivially reflexive and transitive, which, by definition is a partial order.

Next, COORDINATION, or intuitively "all individuals performing the same sequence of actions, at possibly varying delays (Figure 4)," is formally defined as:

Definition 2 (COORDINATION) *Given a set of m-dimensional time series $\mathcal{U} = \{U_1, \ldots, U_n\}$. The set $\mathcal{U}$ is* coordinated *at time t if for every $\binom{n}{2}$ pairs $U_i, U_j \in \mathcal{U}$, either $U_i \prec U_j$ or $U_j \prec U_i$. The* coordination interval *is the maximal contiguous time interval $[t_1, t_2]$ such that $\mathcal{U}$ is coordinated for every $t \in [t_1, t_2]$.*

Finally, the INITIATOR is intuitively "an individual who first performs a sequence of actions, and all other individuals follow," formally defined as:

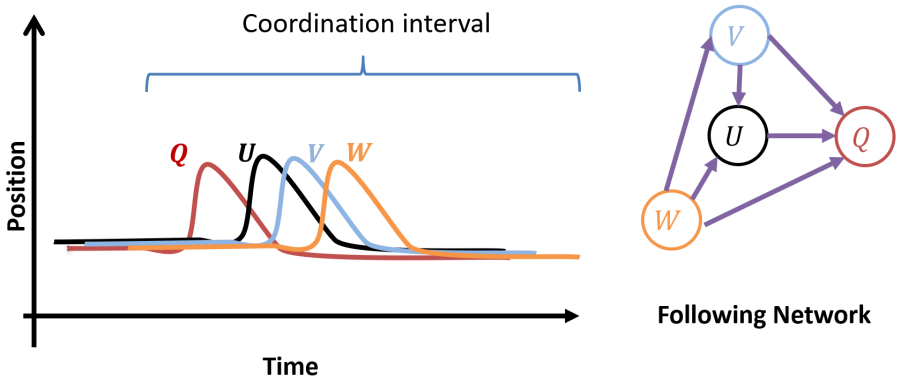


Figure 4. (Left) the example of coordination interval in time series where $Q \prec U \prec V \prec W$. (Right) the following network w.r.t. these following relations. In this example, Q is an initiator.

Definition 3 (INITIATOR) *Let $\mathcal{U} = \{U_1, \ldots, U_n\}$ be a coordinated set of m-dimensional time series within some coordination interval $[t_1, t_2]$. Then the time series $L \in \mathcal{U}$ is the* initiator *time series for the coordination interval if for each time series $U \in \mathcal{U} \setminus \{L\}$, $L \prec U$.*

In Figure 4, Q is an initiator of coordination. The coordination interval starts at the beginning of W. We are now ready to precisely state the problem of identifying the individual who initiates a coordinated behavior:

3.2 Useful observations

Let $\mathcal{U}$ be a coordinated set of time series and $L \in \mathcal{U}$ be the initiator. Since $\mathcal{U}$ is a partial order set and $\forall U_i \in \mathcal{U},\ L \prec U_i$, then, by definition, L is the minimal element. Moreover, $\mathcal{U}$ is a linear order set since for every pair $U_i, U_j \in \mathcal{U}$, either $U_i \prec U_j$ or $U_j \prec U_i$.

Problem 1: COORDINATION INITIATOR INFERENCE PROBLEM

Input : Set $\mathcal{U} = \{U_1, U_2, \ldots, U_n\}$ of time series.

Output: A coordination interval $[t_1, t_2]$ and the initiator time series $L \in \mathcal{U}$ that initiated the coordination.

Definition 4 (FOLLOWING NETWORK) *Let $\mathcal{U} = \{U_1, \ldots, U_n\}$ be a set of time series. The* following network $G = (V, E)$ *is defined as a directed graph where the set of nodes V has a one-to-one correspondence to the set of time series $\mathcal{U}$, and each edge in E represents a* FOLLOWING RELATION *between two time series: $\forall U_i, U_j \in \mathcal{U}$ the edge $e_{i,j} \in E$ if $U_j \prec U_i$.*

Recall that PageRank [66] score, π_i, of a node i in a network G is defined as follows:

$$\pi_i = d \sum_{k \in \mathcal{N}_i^{in}} e_{k,i} \pi_k / |\mathcal{N}_k^{out}| + (1 - d) \tag{3.1}$$

Where $\pi_i \in [0, 1]$, $d \in (0, 1]$ is a constant number, $e_{k,i} \in \{0, 1\}$ is one if $e_{k,i} \in E$, $\mathcal{N}_i^{in}$ is a set of neighbor nodes of i such that $k \in \mathcal{N}_i^{in}$ if $e_{k,i} \in E$, and $\mathcal{N}_i^{out}$ is a set of outgoing neighbor nodes of i such that $k \in \mathcal{N}_i^{out}$ if $e_{i,k} \in E$.

Lemma 3.2.1 *Let $G = (V, E)$ be a following network of time series set $\mathcal{U} = \{U_1, \ldots, U_n\}$. If $U_i \preceq U_j$ then $\pi_i \geq \pi_j$.*

Proof By transitivity, if U_j follows U_i then the followers of U_j are also the followers of U_i. Thus, since $\forall k \in \mathcal{N}_j$, $U_j \preceq U_k$ and $U_i \preceq U_j$, then $\mathcal{N}_j^{in} \subseteq \mathcal{N}_i^{in}$. Hence, $\pi_i - \pi_j = d \sum_{k \in \mathcal{N}_i^{in} \setminus \mathcal{N}_j^{in}} e_{k,i} \pi_k / |\mathcal{N}_k^{out}| \geq 0$.

As a corollary of Lemma 3.2.1, since all the time series follow the initiator L within the coordination interval $[t_1, t_2]$, then L has the highest PageRank score in $\mathcal{U}$ during that coordination period. Moreover, Lemma 3.2.1 allows us to infer the order of following among the time series within the coordination period, as defined by the PageRank values.

3.2.1 Following relation with noise

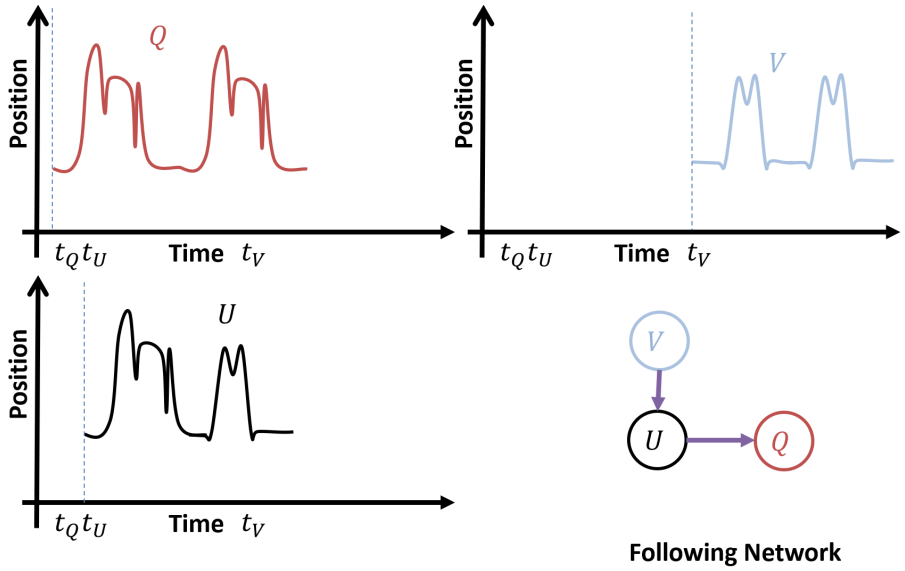


Figure 5. (Top-left) A time series Q, (bottom-left) a time series U, and (top-right) a time series V. In this example, U partially follows Q, V partially follows U but V does not follow Q. The following network of these time series is at the bottom-right of the figure.)

In real situations, Definition 1 requires the exact match, which rarely happens. Therefore, we provide relaxation of following relation to deal with noise in realistic situations below.

Definition 5 (σ-following relation) *Let $\mathcal{U}$ be a set of time series,* $\text{sim} : \mathcal{U} \times \mathcal{U} \rightarrow [0, 1]$ *be a time series similarity function, and $\sigma \in [0, 1]$ be a similarity threshold. For any $P, Q \in \mathcal{U}$, we say that Q σ-*follows *P, denoted as $P \preceq_\sigma Q$, if Q and P are sufficiently similar within some time shift Δ:*

$$\max_{\Delta}(\text{sim}(P, Q_\Delta)) \geq \sigma \text{ and } \min(\underset{\Delta}{\text{argmax}}\, \text{sim}(P, Q_\Delta) \geq 0) \neq \emptyset$$

If either $\Delta t_{max} = 0$ or $U_i \prec_\sigma U_j$ and $U_j \prec_\sigma U_i$, then $U_i \equiv_\sigma U_j$.

The difference between a following relation in Def 1 and Def 5 is that the notion of σ-following relation lacks the transitivity property[1]. Figure 5 shows the example of three time series and their σ-following relation with some unknown $\sigma > 0.5$. In this example, Q is similar to U and U is similar to V greater than 0.5. However, Q and V are not similar at all. Therefore, the σ following relation does not possess the transitivity property.

Next, we can define a notion of coordination by using σ-following relation as follows.

---

[1]This is similar to non-transitive dice: `https://en.wikipedia.org/wiki/Nontransitive_dice`

Definition 6 (*σ*-Coordinated set) *Given a set of m-dimensional time series* $\mathcal{U} = \{U_1, \ldots, U_n\}$ *and a similarity threshold* $\sigma \in (0, 1]$. *The set* $\mathcal{U}$ *is* σ-coordinated *at time t if for every* $\binom{n}{2}$ *pairs* $U_i, U_j \in \mathcal{U}$, *either* $U_i \prec_\sigma U_j$ *or* $U_j \prec_\sigma U_i$.

Definition 7 (*σ*-Coordination interval) *The* σ-coordination interval *is the maximal contiguous time interval* $[t_1, t_2]$ *such that* $\mathcal{U}$ *is coordinated for every* $t \in [t_1, t_2]$.

Even though an initiator of σ-Coordination interval is not a minimum element anymore due to σ-following relation does not possess transitivity property, an initiator suppose to have a highest number of followers. Hence, PageRank is still an appropriate measure for finding an initiator.

3.3 Multiple coordination events without noise

We now extend these concepts to the case of multiple coordinated subgroups.

Definition 8 (Faction) *Given a set of time series* $\mathcal{U}$, *a subset* $F \subseteq \mathcal{U}$ *at time t is maximally coordinated, if F is coordinated and there is no other coordinated set* $F' \subseteq \mathcal{U}$ *where* $F \subset F'$. *We call such maximally coordinated F a faction at time t.*

Definition 9 (Faction interval) *The* coordination interval *of a faction F or a faction interval is the maximal consecutive time interval* $[t_1, t_2]$ *such that F is coordinated for every* $t \in [t_1, t_2]$.

Faction is a structurally maximal subset and its interval is a temporally maximal subset.

Lemma 3.3.1 *A time series W is a member of a faction F if and only if it has an edge to F's initiator L.*

Proof Let a time series $W \in F$. Since $\forall U \in F \setminus \{L\}$, $L \prec W$. By definition, there is an edge from W to L.

Let $L \prec W$. If W is not in F, then we can add W to F, which will remain a coordinated set but will now violate the maximality of F. Thus, $W \in F$.

According to Lemma 3.3.1, a faction F is a set of nodes within G such that all nodes within F have a directed edge to L. Note that L always has the out-degree of zero and in-degree of $|F| - 1$ within a coordination interval. We are now ready to formally state the Faction Initiator Inference Problem at Problem 2.

Problem 2: Faction Initiator Inference Problem

Input : Set $\mathcal{U} = \{U_1, \ldots, U_n\}$ of m-dimensional time series

Output: A set of factions $\mathcal{F} = \{F_1, \ldots, F_k\}$, a set of coordinated intervals

$\mathcal{T} = \{[t_1^1, t_2^1], ..., [t_1^k, t_2^k]\}$, and the set of initiator time series $\mathcal{L} = \{L_1, ...L_k\}$

where L_i initiated the coordination interval $[t_1^i, t_2^i]$ of the faction F_i

3.4 Multiple coordination events with noise

In the previous section, we stated the definitions and properties of the problem of identifying multiple faction initiators in the ideal setting. In this section, we provide the relaxation and the analysis of the problem in the presence of noise.

Definition 10 (σ-faction) *Let $\mathcal{U}$ be a set of time series. A σ-faction $F \subseteq \mathcal{U}$ is a maximal set such that F is σ-coordinated, and there is no other σ-coordinated set $F' \subseteq \mathcal{U}$ where $F \subset F'$.*

Definition 11 (Relaxed faction interval) *Let $\mathcal{U}$ be a set of time series, the time interval $[t_1, t_2]$ is a faction interval of initiator L if for all $t \in [t_1, t_2]$, there exists a faction F_t such that F_t has L as its initiator and $|F_t| > 1$.*

3.4.1 Coordination measure

Given a set of time series $\mathcal{U}$, a set of clusters $\mathcal{C} = \{H_1, \ldots, H_n\}$ such that $\bigcup_k H_k = \mathcal{U}$, we define a cluster membership indicator $\delta_{i,j} = 1$ if time series U_i and U_j belong to the similar cluster, otherwise it is zero. **The average coordination measure** Ψ of a set of clusters $\mathcal{C}$ is defined as follows:

$$\Psi(\mathcal{C}) = \frac{\sum\limits_{U_i, U_j \in \mathcal{U}, U_i \neq U_j} \text{sim}_{max}(U_i, U_j)\delta_{i,j}}{\sum\limits_{U_i, U_j \in \mathcal{U}, U_i \neq U_j} \delta_{i,j}}. \tag{3.2}$$

Note that $\Psi \in [0, 1]$. If Ψ is close to 1, then all time series within the same cluster are highly similar, with some time delay. This implies a high degree of coordination within each cluster in this case. On the contrary, $\Psi \approx 0$ implies no coordination, on average.

Theorem 3.4.1 *Given a set of time series $\mathcal{U}$ containing a set of σ-faction $\mathcal{F} = \{F_1, \ldots, F_n\}$ where $\bigcup_{F_i \in \mathcal{F}} F_i = \mathcal{U}$, then, for all possible sets of clusters, $\mathcal{F}$ maximizes the average coordination measure Ψ.*

Proof Reminding that for all pairs U_i, U_j within any similar faction F, $\text{sim}_{max}(U_i, U_j) \geq \sigma$. Hence, $\Psi(\mathcal{F}) \geq \sigma$.

Case 1: let $H, J \in \mathcal{F}$, if we modify $\mathcal{F}$ by exchanging any time series $U_H \in H$ with $U_J \in J$ and call it $\mathcal{C}$, then we have:

$$\Psi(\mathcal{F}) - \Psi(\mathcal{C}) = \frac{S + S'}{\sum_{U_i, U_j \in \mathcal{U}, U_i \neq U_j} \delta_{i,j}}.$$

$$S = \sum_{U_i \in H \setminus \{U_H\}} \Big(\text{sim}_{max}(U_i, U_H) - \text{sim}_{max}(U_i, U_J) \Big)$$
$$S' = \sum_{U_i \in J \setminus \{U_J\}} \Big(\text{sim}_{max}(U_i, U_J) - \text{sim}_{max}(U_i, U_H) \Big)$$

For any $U_i \in H$, $\text{sim}_{max}(U_i, U_H) \geq \sigma$ since $U_H \in H$. In contrast, because $U_J \notin H$, then $\text{sim}_{max}(U_i, U_J) < \sigma$, which implies $S > 0$. $S' > 0$ for a similar reason. Therefore, $\Psi(\mathcal{F}) - \Psi(\mathcal{C}) > 0$.

Case 2: if we create $\mathcal{C}$ from $\mathcal{F}$ by spiting a cluster $H \in \mathcal{F}$ to be $H_1 \subset H$ and $H_2 = H \setminus H_1$, then we have:

$$\Psi(\mathcal{F}) - \Psi(\mathcal{C}) = \frac{\sum_{U_i \in H_1, U_j \in H_2} \text{sim}_{max}(U_i, U_j)}{|H_1||H_2|} \geq \sigma.$$

Case 3: we create $\mathcal{C}$ from $\mathcal{F}$ by merging any cluster $H \in \mathcal{F}$ with any $J \in \mathcal{F}$ such that $H \neq J$ to be H'. So, let

$$\Psi(\mathcal{F}) = \frac{X_{\mathcal{F}}}{S_{\mathcal{F}}} \geq \sigma,$$

then

$$\Psi(\mathcal{C}) = \frac{X_{\mathcal{F}} + \sum_{U_i \in H, U_j \in J} \text{sim}_{max}(U_i, U_j)}{S_{\mathcal{F}} + |H||J|}.$$

By merging H and J, we introduce pairs of time series across H and J to Equation Equation 3.2 such that $\text{sim}_{max}(U_i, U_j) < \sigma$ since these pairs are not belong to the same faction. These pairs decrease the average of $X_{\mathcal{F}}$, which implies $\Psi(\mathcal{F}) > \Psi(\mathcal{C})$.

Since we shown that no matter how we edit $\mathcal{F}$, the average coordination measure Ψ cannot increase, therefore, $\mathcal{F}$ maximizes the average coordination measure.

3.5 Coordination mechanism inference

We use the following notation throughout this section and Chapter **??**:

- $\mathcal{N} = \{1, \ldots, n\}$ is a set of agents.
- $\mathcal{I} \subseteq \mathcal{N}$ is a set of informed agents.

- $S_i(t)$ is a state value of agent i at time t, where $S_i(t) \in \mathbb{R}^d$.
- $S^t = \{S_i(t)\}$ is a set of individual states at at time t.
- $S_i = (S_i(0), \ldots, S_i(T))$ is state time series of agent i.
- $S_w = (S_w(0), \ldots, S_w(T))$ is a target path where $S_w(t) \in \mathbb{R}^d$ is a target state at time t.
- $\mathcal{H} = \{h_i\}$ is a set of strategy functions that agents use to update their current state where $h_i : \mathbb{R}^d \to \mathbb{R}^d$.
- $\mathcal{S} = \{S_i\}$ is a set of state time series generated by agents using some set of strategy functions $\mathcal{F} \subseteq \mathcal{H}$.
- $\sigma \in [0, 1]$ is a noise-tolerance threshold.

Given a set of n agents $\mathcal{N}$ with a set of their initial states $S^0 = \{S_i(0)\}$, these n agents generate a set of state time series $\mathcal{S} = \{S_i\}$, where $S_i = (S_i(0), \ldots, S_i(T))$ is the state time series of agent $i \in \mathcal{N}$. For each time step t, each agent i updates its state via a strategy function $h_i \in \mathcal{H}$: $S_i(t) = h_i(S_i(t-1))$. However, an informed agent $j \in \mathcal{I}$ always has its state the same as a target path S_w: $S_j(t) = S_w(t)$.

Definition 12 (Coordination event) *Let $\mathcal{Q} = \{Q_1, \ldots, Q_n\}$ be a set of time series. If there exists any σ-coordination interval (Def. 7) in $\mathcal{Q}$, then $\mathcal{Q}$ is a coordination event.*

Definition 13 (Coordination strategy) *Let $\mathcal{F} \subseteq \mathcal{H}$ be a set of strategy functions that the agents use to generate a set of state time series $\mathcal{S} = \{S_i\}$. Each agent $i \in \mathcal{N}$ uses a function $f_i \in \mathcal{F}$ to update its state for each time step. $\mathcal{F}$ is a set of coordination strategies of $\mathcal{S}$ if $\mathcal{S}$ is a coordination event.*

Note that if all agents follow the target path S_w, then an informed agent is an initiator of coordination.

3.5.1 Problem formalization

Suppose there is a set of state time series $\mathcal{S} = \{S_k\}$ that was generated by an unknown set of latent coordination strategies $\mathcal{F} \subseteq \mathcal{H}$ w.r.t. some unknown σ. The only available inputs are $\mathcal{S}$ and the entire set $\mathcal{H}$. The goal is to find $\mathcal{F}$. The real identity of the target path S_w is unknown, but it is known that $S_w \in \mathcal{S}$. Before formalizing the problem, we define the risk function to measure the fitness of any $h_k \in \mathcal{H}$ that might be in $\mathcal{F}$, for any agent i:

$$risk(S_i, h_k) = \frac{1}{T} \sum_{t=1}^{T} loss(S_i(t), h_k(S_i(t-1))), \tag{3.3}$$

where $loss : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a loss function and $h_k(S_i(t-1))$ returns a predicted state $\hat{S}_i(t)$. Now, we are ready to formalize Coordination Strategy Inference Problem.

Problem 3: Coordination Strategy Inference Problem

Input : A set of state time series $\mathcal{S} = \{S_i\}$ generated by multiple agents, where $\mathcal{S}$ is a coordination event; a set of strategy functions $\mathcal{H} = \{h_k\}$; and a loss function $loss : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$.

Output: A set of minimum risk strategies $\mathcal{F}^* = \{f_i^*\}$ where, for each agent i, $f_i^* = \text{argmin}_{h_k \in \mathcal{H}}\, risk(S_i, h_k)$.

3.6 Mining patterns of leadership dynamics

Let $\mathcal{U}$ be a set of time series, We create a dynamic following network $\mathcal{G} = \langle G_t \rangle$ by considering each temporal sub-interval t of $\mathcal{U}$ of length ω (time window parameter) and creating a following network G_t of that interval. We then define the notion of the time series of leaders of a dynamic following network below.

Definition 14 (Time series of leaders) *Let $\mathcal{U}$ be a set of time series. $\mathcal{L}$ is a* time series of leaders *where $\mathcal{L}(t)$ is a set of faction initiators at time t in G_t.*

We can use mFLICA framework [4] to extract a time series of leaders from time series of movement. Next, we define the support of a leader set S. Let T be the length of the time series of leaders and $\mathbb{1}_x$ be an indicator function, which is 1 if the statement x is true, and 0 otherwise.

$$\text{supp}_{\mathcal{L}}(S) = \frac{\sum_{t=1}^{T} \mathbb{1}_{S=\mathcal{L}(t)}}{T}. \tag{3.4}$$

$\text{supp}_{\mathcal{L}}(S)$ indicates the support of having a particular set of initiators S lead multiple groups concurrently. For example, if $\text{supp}_{\mathcal{L}}(\{L_1, L_2\}) = 0.5$ it means that half the time the leaders are exactly $\{L_1, L_2\}$, leading their factions concurrently.

Definition 15 (Frequent-leader set) *Let $\mathcal{L}$ be a time series of leaders, S be a set of faction initiators, and $\phi \in [0, 1]$ be a support threshold. S is a frequent-leader set of $\mathcal{L}$ if* $\text{supp}_{\mathcal{L}}(S) \geq \phi$.

Definition 16 (Transition probability of leader sets) *Let $\mathcal{L}$ be a time series of leaders, and S_i, S_j be sets of faction initiators. A transition probability of leader sets λ_{S_i,S_j} is a probability that $\mathcal{L}(t-1) = S_i$ and $\mathcal{L}(t) = S_j$.*

Now, we are ready to formally state the the problem of Mining Patterns of Leadership Dynamics.

Problem 4: Mining Patterns of Leadership Dynamics

Input : A set $\mathcal{U} = \{U_1, \ldots, U_n\}$ of m-dimensional time series and a support threshold ϕ.

Output: A set of frequent-leader sets $\mathcal{S}_\mathcal{L}$ and a transition probability set $\mathcal{P} = \{\lambda_{S_i,S_j}\}$ where $S_i, S_j \in \mathcal{S}_\mathcal{L}$.

In this work, we choose to represent a set of frequent-leader sets as a diagram of leadership dynamics below.

Definition 17 (A diagram of leadership dynamics) *Let $\mathcal{L}$ be a time series of leaders, $\phi \in [0, 1]$ be a support threshold, and $S_\mathcal{L}$ be set of frequent-leader sets. A digraph $\mathcal{T} = (V_\mathcal{T}, E_\mathcal{T})$ is a diagram of leadership dynamics such that the nodes $V_\mathcal{T}$ represent frequent-leader sets $S_\mathcal{L}$ and $(v_i, v_j) \in E_\mathcal{T}$ if $\lambda_{S_i,S_j} > 0$.*

CHAPTER 4

LEADERSHIP MODELS AND DATASETS

In this chapter, we provide the details of leadership models in simulated datasets that we use to evaluate our framework performance as well as the details of three real-world datasets: trajectories of baboons, trajectories of schools of fish, and time series of stock closing-price from NASDAQ.

4.1 Simulation of leadership models

4.1.1 Dictatorship model (DM)

In this model, we fix a single initiator who initiates movement from initial positions of the population. At the start of the pre-coordination interval, the initiator moves in a fixed direction and acceleration. Other individuals wait for a randomly sampled lag, before following the initiator at a fixed acceleration (with sampled noise in the heading). After a fixed duration of coordinated movement over the entire population, individuals decelerate at random, until stopping. Figure 6 (left) shows the example of DM. The Switching Dictatorship model (DM-S) selects two fixed individuals over each trial: a single individual as an initiator during pre-coordination, and another single individual as 'initiator' during coordination. Figure 7 shows an example of following-network-density time series of DM-S.

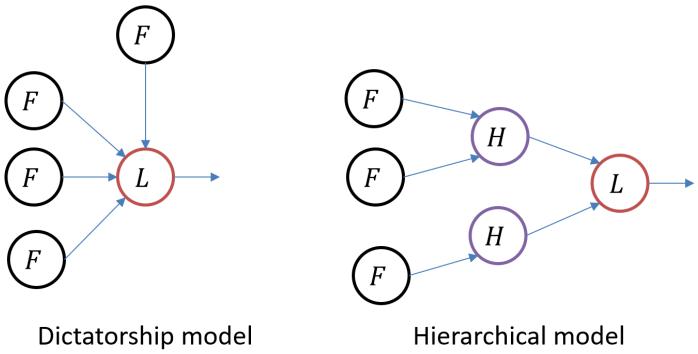


Figure 6. Examples of individual movements from Dictatorship and Hierarchical model. Nodes represent individual positions and arrows represent directions of individual's movement. (Left) in Dictatorship model, everyone follows a leader L, while there is a hierarchy to follow for each individuals in Hierarchical model (right).

4.1.2 Hierarchical model (HM)

This model is a variation of DM, where we fix a number of individuals (n=4) to follow the previous individual in the sequence, after a sampled lag. The remainder of individuals in the population follow exactly one of these high-ranking individuals, allocated in decreasing proportion per rank. Figure 6 (right) shows the example of HM. The Switching Hierarchical model (HM-S), similarly to DM-S, selects unique pairs of individuals for each hierarchy level, switching after the pre-coordination interval as in DM-S.

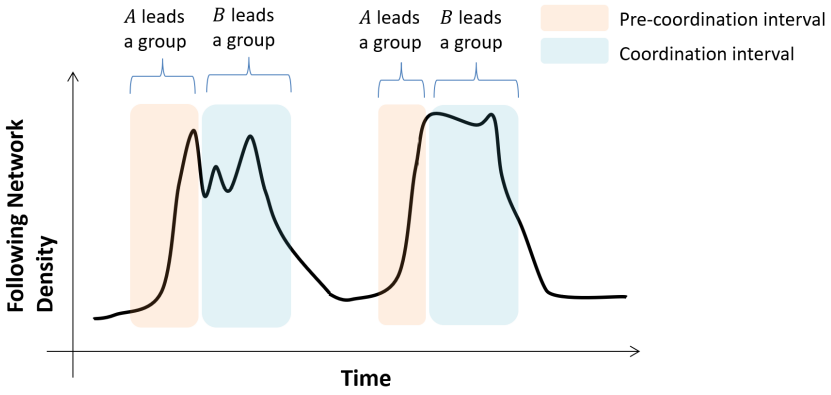


Figure 7. An example of following-network-density time series of Switching Dictatorship model . There are two coordination events where individual A leads a group in both pre-coordination intervals while B leads a group during coordination intervals.

4.1.3 Event-based model (EM)

This model is a variation of the Dictatorship model where each coordination event has a different, unique initiator. For example, in one of our applications, a troop of baboons may follow an initiator to a food source in the morning, and follow a different initiator in the evening to the sleeping site. No existing methods can infer these two situations except our framework. Figure 8 shows an example of following-network-density time series of EM.

4.1.4 Initiator model (INIT-k)

In this model, we fix k initiators who initiate movement from random initial positions of the population. At the start of the pre-coordination interval, all initiators move on a single target. Non-initiators move in randomly sampled directions with a fix velocity, then follow

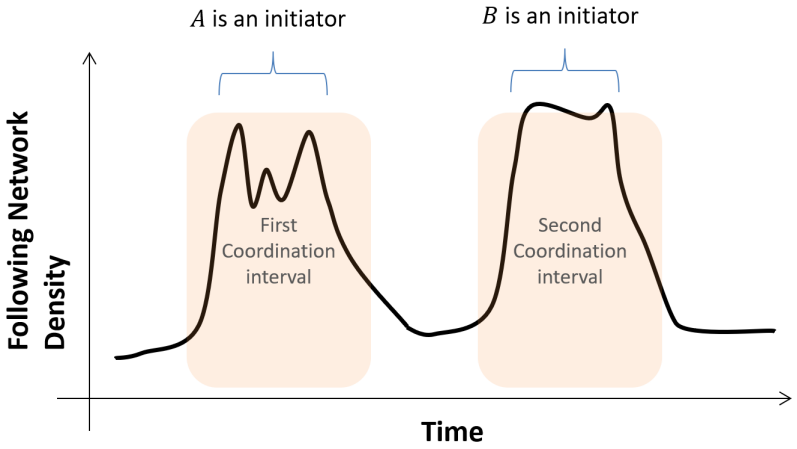


Figure 8. An example of following-network-density time series of Event-based model. There are two coordination events where individual A leads a group in the first coordination event while B leads a group during the second coordination event.

their initiators after a random time lag. After the pre-coordination period, all individuals move toward a single target, without following their initiators. The example of INIT-k model is at Figure 9. We run simulations for INIT-1 and INIT-4 initiator models.

4.1.5 Crowd model (CM)

This model [25] is a collective movement model where k (=4) informed individuals move toward a target, and the remaining (=16) uninformed individuals move in a linear combination of a direction toward the group's centroid, and the average direction of the group. The example of CM is at Figure 10.

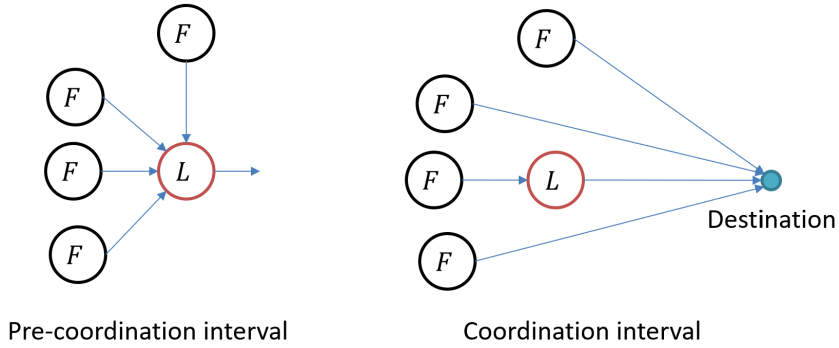


Figure 9. Examples of individual movements from Initiator model. Nodes represent individual positions and arrows represent directions of individual's movement. (Left) during a pre-coordination interval, everyone follows a leader L, while, during a coordination interval, everyone knows the direction and moves to the destination directly without following a leader (right).

4.1.6 Linear Threshold model (LT)

This model [36] initiates individual movement by propagation of a linear threshold process on the dynamic network, defined by the k-nearest neighbors at the current time-step. The model is parameterized by ρ, the proportion of these k neighbors required to be infected in order to initiate movement. Once activated, the individual follows a single initiator. The initial probability of activation for each individual is 0.5. We explore the parameter space on combinations of: $k \in \{3, 5, 10\}$ and $\rho \in \{0.25, 0.50, 0.75\}$. The example of LT is at Figure 11.

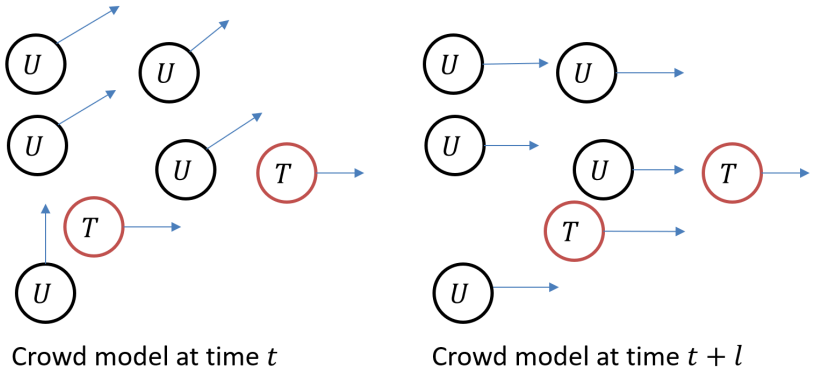


Figure 10. Examples of individual movements from Crowd model. Nodes represent individual positions and arrows represent directions of individual's movement. (Left) at time t, everyone follows some directions except informed individuals (T nodes) which moves directly to a target. Then, at time $t + l$ (right), the group's direction, which is the average of individual's directions, gradually changes toward the target.

4.1.7 Independent Cascade model (IC)

This model [36] is another propagation process similar to LT. At each time step, each active individual moves toward the initiator and independently attempts to activate its k-nearest neighbors with the probability of ρ. If the individual fails to activate a neighbor, it cannot attempt to activate the same neighbor again. We explore the same sample parameter space as in the LT model.

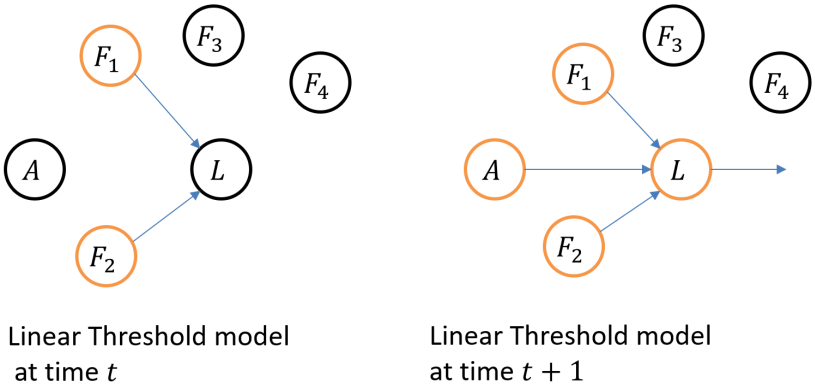


Figure 11. Examples of individual movements from Linear Threshold model. Nodes represent individual positions and arrows represent directions of individual's movement. (Left) at time t, only active individuals (orange color) move toward an initiator L. Suppose $k = 3$ and $\rho = 0.50$, at time t, an inactive individual A has two active individuals F_1, F_2 and L as its neighbors. Since 66% of A's neighbors are activated, then, at time $t + 1$ (right), A is active and start moving toward L.

4.1.8 Random model

In this model, there is no 'following' relations. At the start of the pre-coordination interval, all individuals start moving to a fixed direction, independently of others in the population. We expect the relative positions of individuals to yield some following relations only by chance.

4.2 Real-world datasets

4.2.1 Baboon trajectories

High-resolution GPS collars track 26 individuals of a troop of olive baboons (*Papio anubis*) living in the wild in Mpala Research Centre, Kenya [6, 7]. The data consists of latitude-longitude location pairs for each individual at one observation per second. We analyze a subset of 16 individuals whose collars remained functional for a ten day period (419,095 time steps). In addition, in the first two days of baboon tracking, there are four group activities labeling by experts: sleeping, hanging out, coordinated progression, and coordinated non-progression [67]. We show later that by using only following network density as a feature to perform activity classification, we can get high accurate results of activity prediction.

4.2.2 Fish schools trajectories

The movement of a fish school of golden shiners (*Notemigonus crysoleucas*) are recorded by video in order to study information propagation over the visual fields of fish [46]. Each population contains 70 fish, with 10 trained, labeled fish who are able to lead the school to feeding sites over 24 separate coordination events. The task is to correctly identify trained fish by initiator ranking.

4.2.3 Stock closing-price time series

We collected daily closing price data for stocks listed in NASDAQ, using Yahoo! Finance.[1] These time series are from January 2000 to January 2016 (4169 time-steps). We remove symbols

[1] http://finance.yahoo.com/

with a large amount of missing data, leaving a total of 1443 symbols in our dataset. Our analysis focuses on discovering large, known events and crises in an unsupervised way, and to explore initiators and sectors involved in these coordination events.

CHAPTER 5

FLICA: A FRAMEWORK FOR LEADER IDENTIFICATION IN COORDINATED ACTIVITY

COORDINATION INITIATOR INFERENCE PROBLEM: An agreement of a group to follow a common purpose is manifested by its coalescence into a *coordinated* behavior. The process of initiating this behavior and the period of decision-making by the group members necessarily precedes the coordinated behavior. **Given time series of group members' behavior, the goal is to find these periods of decision-making and identify the initiating individual, if one exists.**

5.1 Introduction

In this chapter, we present a Framework for Leader Identification in Coordinated Activity (FLICA) as the solution for the COORDINATION INITIATOR INFERENCE PROBLEM. On real data, the formalization in Chapter 3 is very restrictive, so we relax the *exact* FOLLOWING RELATION, and *full* COORDINATION to identify 'following' and partial 'coordination' in real

---

This chapter has been previously published in Chainarong Amornbunchornvej, Ivan Brugere, Ariana Strandburg-Peshkin, Damien R. Farine, Margaret C.Crofoot, and Tanya Y. Berger-Wolf. 2018. Coordination Event Detection and Initiator Identification in Time Series Data. ACM Trans. Knowl. Discov. Data.12, 5, Article 53 (June 2018) https://doi.org/10.1145/3201406 See Appendix A for the copyright permission document from the publisher.

applications. Furthermore, multiple coordination events often exist within a set of real time series data. Constructing a single aggregated network would not capture the dynamics of these events. Therefore, FLICA uses a dynamic network approach.

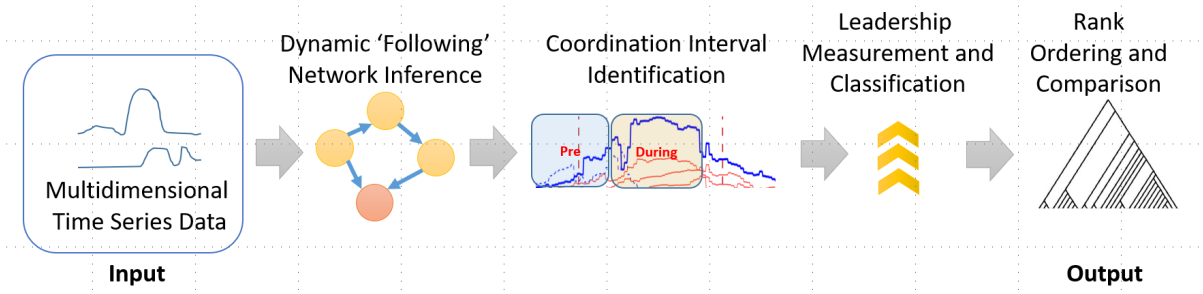
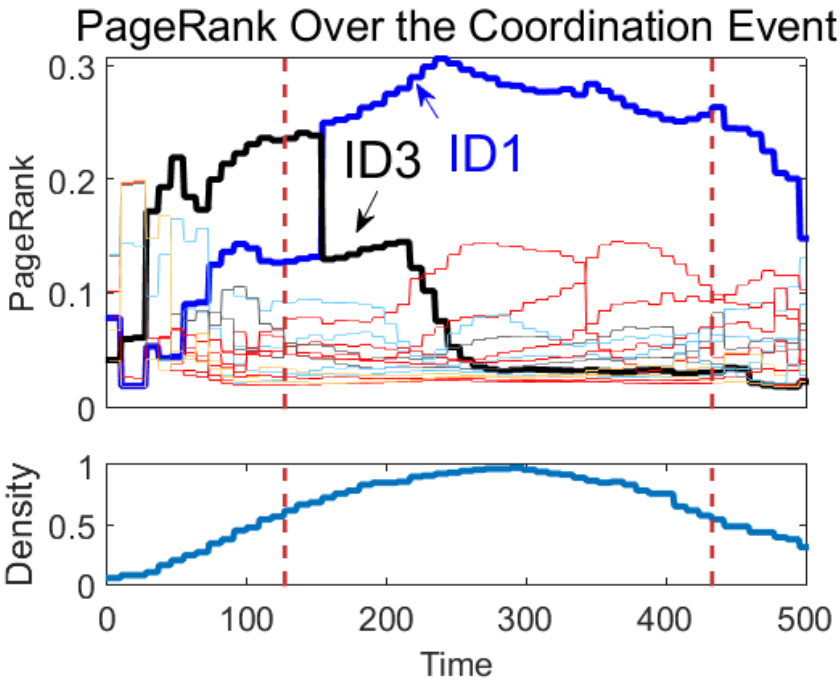


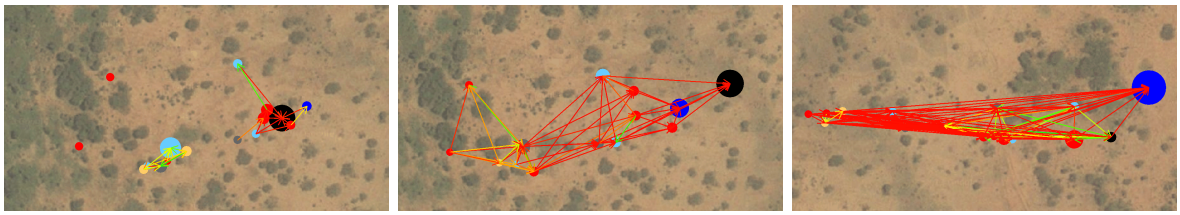
Figure 12. A high-level overview of FLICA framework

Figure 12 shows the framework overview. At each time step, we infer following relations to construct a sequence of following networks. We then use network density to detect intervals of coordination, and the time series of PageRank values to identify the initiators of these coordination intervals.

5.1.1 A working example

Figure 13 presents a key example and a brief introduction to our framework, on real GPS trajectory data of olive baboons (*Papio anubis*). Figure 13(b) and Figure 13(c) show the leadership of movement of the group by baboon ID3 (Black). Figure 13(d) shows the 'following'





(b) t=50 (c) t=100 (d) t=250

Figure 13. PageRank (top) and density (middle) of the 'following' network over time for an event of baboons' movement which initiates by ID3. (Bottom) The locations of individuals over three different time steps ($t = 50, 100, 250$), with the 'following' network, and PageRank indicated by node size.

network in the coordination interval. Individual ID3 has the largest PageRank in the first two snapshots but the PageRank of individual ID1 (Blue) surpasses ID3 when the network is 'coordinated' (e.g. moving together). If we measure the initiator ranking *after* the network has coalesced, then we miss that ID3 initiated coordination and 'built' the network in the pre-coordination interval (to the left of the first dotted red line).

We now present each step of the computational framework of FLICA. We will discuss the following relation inference in Section 5.2.1, the construction of the following network in Section 5.2.2, identification of the coordination interval and the preceding decision-making period in Section 5.2.3, the identification of the initiator in Section 5.2.4 and the details of model and parameter choices at each step.

5.2 Methods

5.2.1 Following relation inference

Given a pair of time series U, Q, our task here is to find a following relation between U and Q. However, we relax a notion of following relation in Def. 1 to allow some degree of distortion between two time series that follows each other. A measures we need should satisfy two properties. First, a measure must be able to identify common pattern between U and Q. A common pattern might not happens in Q the same time as U and the common pattern might have some degree of distortion. Second, a measure must be able to infer a time delay between common patterns in U and Q. With these properties, for a (U, Q) pair of time series, we use Dynamic Time Warping (DTW) [68] to measure whether U follows Q. DTW is shown

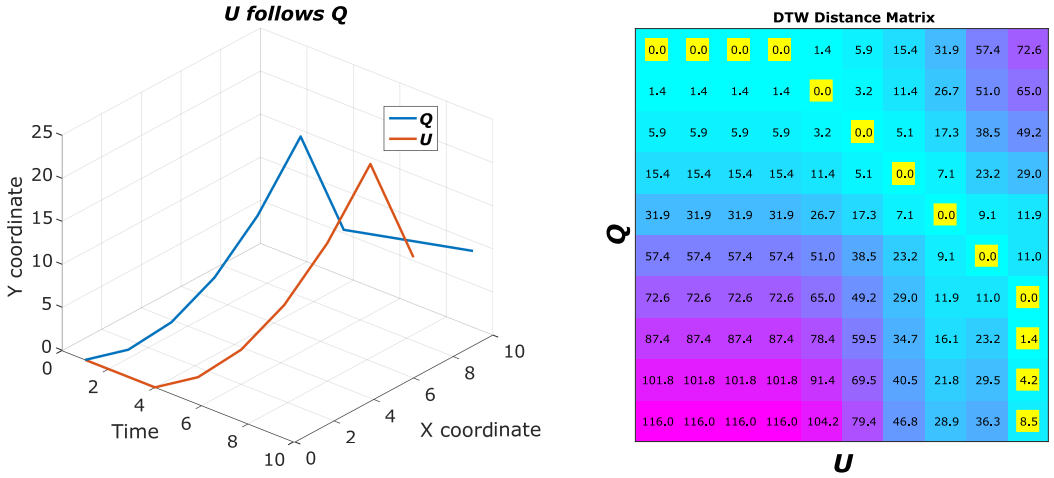


Figure 14. (Left) Toy time series showing U following Q with a time delay $\Delta = 3$. (Right) the optimal warping path (yellow boxes) on the DTW dynamic programming matrix, shifting U backward in time onto Q.

to perform better than several other methods in inferring following relation in time series [33] and it is tolerant to noise [69]. Figure 14 (Left) shows two time series, where time-shifting Q ahead in time produces a better match to U, illustrated in the *warping path* in Figure 14 (Right). Let $P_{U,Q}$ be a sequence of index pairs (i, j) which comprise the DTW optimal warping path of (U, Q). We compute the mean of the signed index difference over this sequence of index pairs:

$$\mathrm{s}(P_{U,Q}) = \frac{\sum_{(i,j)\in P_{U,Q}} \mathrm{sign}(j-i)}{|P_{U,Q}|} \tag{5.1}$$

This function measures the extent of warping between two time series. If time series cannot be shifted one-onto-the-other with a consistent positive or negative sign, $|\mathrm{s}(P_{U,Q})| \approx 0$, then there is no following relation between U and Q. When $\mathrm{s}(P_{U,Q})$ is positive, Q *follows* U, otherwise, U *follows* Q. In Figure 14, $\mathrm{s}(P_{U,Q}) \approx -3$.

5.2.2 Dynamic following network inference

As shown in Section 5.1.1, a coordinated activity is dynamic in the aspect of who leads a group at each time step. Using only summary statistics of static following network to represent the entire coordinated activity cannot capture dynamics of coordinated activity. Therefore, we deploy a dynamic network procedure to analyze coordinated activities in time series, which is a common technique to deal with dynamics of data [70].

In our setting, the set of n m-multidimensional time series $\mathcal{D}$ (*e.g.,* a matrix of size $[n \times m \times t^*]$), a window size parameter ω, and a window shift parameter δ (default is 0.1ω) are the inputs for our framework.

Let the ith time interval be given by: $w(i) = [(i-1) \times \delta, (i-1) \times \delta + \omega]$. For each $w(i)$, we extract a set of sub time series $\mathcal{Q}_i$ from $\mathcal{D}$. The $\mathcal{Q}_i$ is the $[n \times m \times \omega]$ dimensional matrix of the time series set. Then we construct a following network $G = (V, E)$ as defined in Def. 4. The nodes represent the time series from $\mathcal{Q}_i$ and E is a set of edges between time series nodes such that if $U, W \in \mathcal{Q}_i$ and U follows W according to Eq. Equation 9.1, then $e_{U,W} \in E$ with the edge weight $|\mathrm{s}(P_{U,W})|$. We calculate a following network for each $w(i)$ to construct a dynamic following network $G^* = (V, E^*)$. The pseudo code is given in Procedure 5.

Procedure 5: CreateDyFollowingNetwork

input : A set of time series $\mathcal{D}$, a time window ω, and a window shift δ

output: An $n \times n \times t^*$ adjacency matrix E^*.

$K \leftarrow (t^* - \omega)/\delta$;

for $i \leftarrow 1$ **to** K **do**

/* current time interval */

$w(i) = [(i-1) \times \delta, (i-1) \times \delta + \omega]$;

/* SubTimeSeries($\mathcal{D}, w(i)$) returns all sub time series in $\mathcal{D}$ within the interval $w(i)$ */

$\mathcal{Q}_i \leftarrow$SubTimeSeries$(\mathcal{D}, w(i))$;

$E \leftarrow$CreateFollowingNetwork$(\mathcal{Q}_i)$;

/* Set all edges within the time interval $[(i-1) \times \delta, i \times \delta]$ to be similar */

$E^*_{t \in [(i-1) \times \delta, i \times \delta]} \leftarrow E$;

end

$\mathcal{Q} \leftarrow$SubTimeSeries$(\mathcal{D}, [K \times \delta, t^*])$;

$E \leftarrow$CreateFollowingNetwork$(\mathcal{Q})$;

$E^*_{t \in [K \times \delta, t^*]} \leftarrow E$;

5.2.3 Coordination intervals detection

Network density of the following network serves as the measure of the extent of coordination over all time series pairs (by Def. 2, during the coordination interval *every* pair has a following relation.) We can use this observation to identify times of *approximate coordination*.

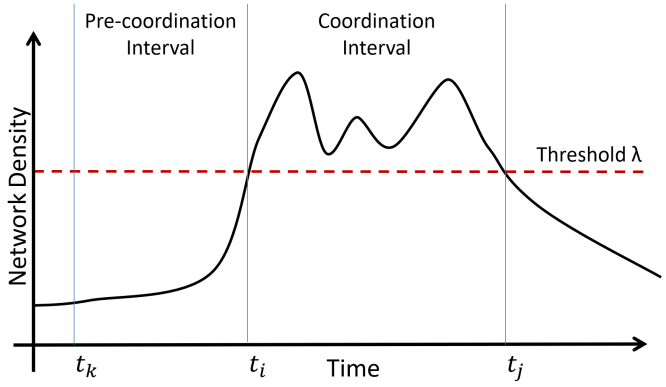


Figure 15. A *coordination event* is a pair of intervals. We define the *pre-coordination interval* and *coordination interval* using threshold λ on the network density time series.

Given a time series of network densities, denoted by d, over a dynamic following network G^*, and a density threshold parameter λ, the time interval $[t_i, t_j]$ is a λ-coordination interval if $\mathrm{d}(t) > \lambda$ for *all* $t \in [t_i, t_j]$. The *pre-coordination interval* of coordination $[t_i, t_j]$ is the interval $[t_k, t_i - 1]$, where the discrete derivative $\mathrm{d}(t) - \mathrm{d}(t-1) \geq 0$ for all $t \in [t_k, t_i - 1]$. Together, these intervals are one *coordination event*, represented by the 3-tuple of time indices $I = (t_k, t_i, t_j)$. The collection of coordination events is a set $C = \{I_l\}$. All complete event intervals $[t_k, t_j]$ are mutually disjoint in C, and $|C|$ denotes the total number of 3-tuples. Figure 15 illustrates the definition of a *coordination event* as a pair of time intervals. To reduce the number of intervals generated near the threshold λ, we apply a greedy merging of nearby coordination intervals (taking the range from the window size ω).

5.2.4 Ranking comparison

On each coordination event $I = (t_k, t_i, t_j)$, let R_I be some ranking of individuals within the pre-coordination interval $[t_k, t_i - 1]$. We focus on ranking within pre-coordination because this is the interval where coordination is initiated. The *global* rank order of pre-coordination, denoted by $\hat{R}$, is the average of all R_I where $I \in C$.

We measure **initiator ranking** according to three different methods: PageRank [66], velocity convex hull (VCH), and position convex hull (PCH). Recall, that by Lemma 3.2.1, if U follows V, then the PageRank of U is less than that of V. Thus, the initiator is expected to have the highest PageRank. VCH measures how often an individual moves faster than others. It represents a model of leadership for movement. This model can be found in many social species [8, 23]. PCH measures how often an individual moves to an area before others. For example, in a flock model [30], a leader is positioned at the front of the group's trajectory.

5.2.4.1 PageRank

PageRank is a standard method for measuring the importance of a node recursively by the importance of the nodes linking to it. In a directed network where a link represents a *following* relations between nodes, PageRank measures 'following' *paths* passing through a particular node. Thus, it fits well with our definition of leadership.

PageRank returns a weight vector of length n, with a sum of 1. For each time step t, we calculate PageRank for each static graph G_t within a dynamic following network $G^* = (V, E^*)$. Let $\mathcal{R} = (R_{pr}(1), \ldots, R_{pr}(t^*))$ be a sequence of n-length PageRank Order vectors where $R_{pr}(t) = \text{argsort}(PageRank(G_t))$ such that $R(i)_{pr}(t)$ represents the rank of individual i

and $R(i)_{pr}(t) < R(j)_{pr}(t)$ if the PageRank value in Eq. Equation 3.1 of i is greater than the value of j ($\pi_i > \pi_j$.) The leader L at time t is the individual who has the highest value of PageRank π_L or $R(L)_{pr}(t) = 1$. Note that argsort$(*)$ returns the index list of sorted values w.r.t. descending order.

5.2.4.2 Velocity Convex Hull

In the next two sections, the s-energy work by Chazelle [47] motivates the use of the convex hull as the measure of the level of initiation of a state change for the group. Chazelle showed that if every agent in a group remains within the convex hull of its neighbors (even if the neighbors change) at each time step, then the system converges to an equilibrium. Thus, to change the steady state, somebody needs to break out of the convex hull of their neighbors. In the initial state, all individuals states such as velocity or position are inside the group's convex hulls of that state. Then after the group decides to change its state, some individuals must step outside the group convex hull to make the change. Hence, by using convex hull analysis, we can measure whether the initiators are also state changers. Specifically, we use convex hull (of position and velocity) analysis to characterize leadership models.

The velocity convex hull measures the frequency with which the discrete time series derivative (dQ/dt) associated with a node i is outside the bounds of the population's discrete derivative distribution (including node i) in the previous time step. In aggregate, a high rank of this measure indicates which node first 'moves' in the group.

The convex hull can be computed on arbitrary m dimensions of a multidimensional time series, or their derivatives, jointly or independently. The convex hull function CH$(*)$ returns an m-dimensional surface represented as lines between points in the input data, which encompass all other points.

Let $\mathbb{V}$ be a $[n \times t-1]$-sized matrix measuring individual velocity over time, on time series dataset $\mathcal{D}$, which is a $[n \times m \times t^*]$-sized matrix. $\mathcal{D}_{i,k}(t)$ represents a data point of kth dimension of time series i at time t. We refer to $\mathcal{D}_{i,*}(t)$ as a m-dimensional data point of time series i at time t[1]. For an individual i at time-step t, we define the following indicator function:

$$\text{VCH}(\mathbb{V}, i, t) = \begin{cases} 1, & \mathbb{V}_i(t) > \max(\mathbb{V}_*(t-1)) \\ -1, & \mathbb{V}_i(t) < \min(\mathbb{V}_*(t-1)) \\ 0, & \text{otherwise} \end{cases} \tag{5.2}$$

For time step j we output an n-length rank order vector as $R_v(t) = \text{argsort}((\text{VCH}(\mathbb{V}, i, t))_{i=1\ldots n})$.

5.2.4.3 Position Convex Hull

The position convex hull is analogous to velocity, except that our indicator function measures an individual's position relative to the convex hull containing the population at the previous time step. Rather than look at velocity of initiation, this measure captures an individual's frequency of moving outside the geometric boundaries of the group in the time series space, and close to the average heading of the group (e.g. in 'front' of the group).

---

[1] We use '*' subscript notation in matrices to indicate slicing in the dimension(s).

We compute the convex hull function on time-step t, $H(t) = \text{CH}(\mathcal{D}_{*,*}(t))$, and also introduce the heading vector of individual i: $\vec{v}_i(t) = (\mathcal{D}_{i,*}(t-1), D_{i,*}(t))$, and the population heading vector: $\vec{v}(t) = (1/n)\sum_{i=1..n} \vec{v}_i(t)$. We define the function IN(A, B) to denote standard 'B contains A' spatial queries between two geometry objects, and $\measuredangle(\vec{v_1}, \vec{v_2})$ to measure the angle between two vectors $\vec{v_1}$ and $\vec{v_2}$.

Using these definitions, we define the position convex hull indicator function for individual i at time t:

$$\text{PCH}(D, i, t) = \begin{cases} 1, & \neg\text{IN}(\mathcal{D}_{i,*}(t), H_{t-1}), \measuredangle(\vec{v}_i(t), \vec{v}(t)) \leq 90^\circ \\ -1, & \neg\text{IN}(\mathcal{D}_{i,*}(t), H_{t-1}), \measuredangle(\vec{v}_i(t), \vec{v}(t)) > 90^\circ \\ 0, & \text{otherwise} \end{cases} \tag{5.3}$$

For time step t we output an n-length rank order vector as $R_p(t) = \text{argsort}((\text{PCH}(D, i, t))_{i=1...n})$.

5.2.5 Leadership model features

Let the global rank ordering of pre-coordination for PageRank be denoted by $\hat{R}_{pr}$, for VCH by $\hat{R}_v$, and for PCH by $\hat{R}_p$. To measure global leadership of pre-coordination in our framework, we order individual nodes i based on initiation support with respect to one of these ranking methods, $R_{*,I}$, over all coordination events $I \in C$. For example, $R_{pr,I}$ is PageRank-rank-ordered list at the coordination event I. If an individual i is at 1st rank at I, then $(i, 1) \in R_{pr,I}$;

i is an initiator. The initiation support for a node i is the fraction of coordination events at which it was ranked 1 (by a ranking measure R_*):

$$\sup_*(i) = \frac{|\{\forall I \in C\,(i,1) \in R_{*,I}\}|}{|C|} \tag{5.4}$$

We use the Kendall rank correlation coefficient $\tau()$ [71] to compare event-local and global rank-orders. To compare global and local rank orders, we use the mean Kendall rank correlation over all coordination events against the global by Eq. Equation 5.5. For example, corr_v compares local and global velocity-convex-hull-rank orders.

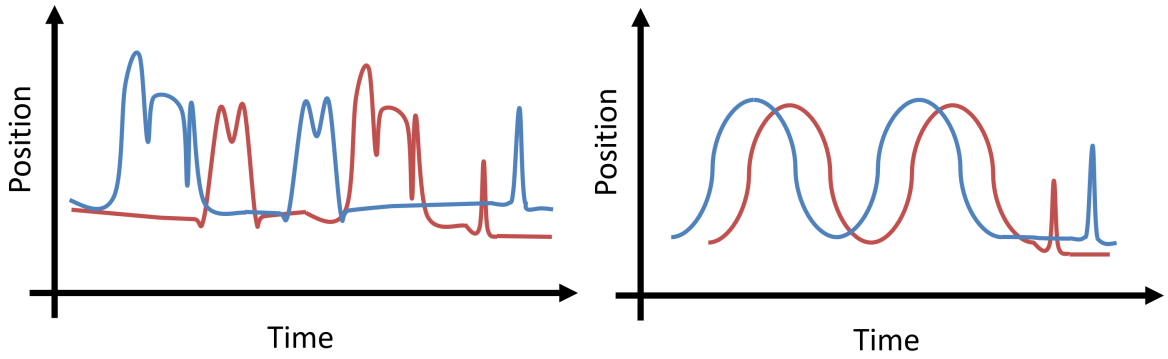
$$\text{corr}_* = \frac{\sum_{I \in C} \tau(\hat{R}_*, R_{*,I})}{|C|} \tag{5.5}$$

Similarly, we compute the mean Kendall correlation between local rankings associated with different measures (e.g. VCH, PCH) by Eq. Equation 5.6.

$$\text{corr}_{*,*} = \frac{\sum_{I \in C} \tau(R_{*,I}, R_{*,I})}{|C|} \tag{5.6}$$

corr_* formalizes our intuition that leaders consistently move outside of the spatial extent (corr_p), or the distribution of velocity over the population (corr_v). By comparing the global vs. local correlation in rank ordering, we measure the stability of the global ranking is over time.

$\text{corr}_{*,*}$ measures the relationship between higher-order graph structure and simple time series features. Using this measure, we can gain a better understanding of the high-level aspects of



(a) Pattern order alignment (b) Global shift alignment

Figure 16. Dynamic Time Warping global vs. local example

initiating coordination. For example, we see whether changing velocity ($\text{corr}_{v,pr}$), or position ($\text{corr}_{p,pr}$) within the group is correlated with network rank position.

5.2.6 Local vs. Global Matching

Our proposed framework uses local alignment on time series subsequences, rather than global alignment on the full time series. Figure 16 presents a motivation for this choice. Suppose we intend to match *sparse* 'following' events represented as the pair of spikes with relatively low magnitude at the end of the red and blue time series. In Figure 16(a), the time series is shifted to match one of the two patterns, depending on the cost. This forces a mismatch of the 'following' event. Similarly, Figure 16(b) has a low cost matching by shifting the entire time series at a constant rate. By matching only local subsequences, we can recover both of these 'following' events.

5.3 Experimental setup

We evaluate our framework on eight synthetic movement trajectory models in Section 4. We also use three real datasets in Section 4.2 to demonstrate the utilities of our framework.

5.3.1 Synthetic trajectory simulation

For each of the above models, we generate a trial of synthetic data consisting of 20 individuals, and 20 separate coordination events, for a total of 12,000 time-steps. Each coordination event has pre-coordination and coordination intervals of 200 time-steps each. Following the coordination interval is another 200 time steps of a post-coordination before repeating. We generate 100 trials for each of models. In total, we have 2,700 simulation datasets.

5.3.2 Evaluation

For synthetic datasets, we use three evaluation approaches:

- **Global leadership**: For each method, we extract network and/or rank statistics over the entire time series, and report only a single aggregate initiator ranking. We compare the known ground truth ranking (used to generate the data) against the ranking of each method, reporting precision. We measure precision of identifying the true initiator, on DM, LT, IC, and INIT-1 models. For the HM model, we compare the *exact* top-4 ranking against the ground truth (order matters); The evaluation is the same for CM and INIT-4 models, except the *exact* top-4 ranking constraint is relaxed (i.e. we compare top-4 *sets*).
- **Local leadership**: For evaluation data in this case, we use the ground truth ranking for each local coordination event, and the time intervals of each event. We report average precision over each discovered pre-coordination interval. We evaluate the EM model using this approach.

We report only the FLICA result, since it is the only method capable of producing local ranking.

- **Initiator leadership**: For each coordination event, we measure the initiator of coordination event before coordination occurs (e.g. in the pre-coordination interval). This individual may not be highly ranked after coordination (see: Figure 13). We report the precision of global leadership considering only the pre-coordination intervals. Since only FLICA identifies pre-coordination intervals, we compare against other methods' global leadership. This evaluation demonstrates that global leadership is distinct from coordination initiation. We evaluate DM-S and HM-S models using this approach.

5.3.3 Compared leadership methods

We demonstrate the performance of our framework by comparing with previous works on influence and leadership [30, 33, 36] as well as creating the Granger-Causality framework based on the work by Liu et al. [72] to illustrate the potential of using Granger Causality to infer leaders in time series. These methods can infer only global initiator ranking, while our proposed framework (FLICA) can detect individual coordination events, handles switching initiator, and performs leadership model classification. Therefore, we use the global leadership identification task to compare FLICA's performance with the prior works. We report the best results under varying parameters for competing frameworks. The time complexity of each method is shown in Table Table I.

First, the FLOCK model [30] identifies leaders who move toward the norm direction vector of the group and also in the front of the group. Second, LPD [33] creates an aggregate 'following'

TABLE I

Time complexities of leadership inference methods where n is a number of time series, ω is a time window, and t^* is a length of time series.

Method	Input	Time complexity
FLICA	Time series	$\mathcal{O}(n^2 \times t^* \times \omega)$
FLOCK [30]	Trajectory	$\mathcal{O}(n^2 \times t^*)$
LPD [33]	Time series	$\mathcal{O}(n^2 \times t^* \times \omega^3)$
IM [36]	Network	$\mathcal{O}(n^2 \times t^*)$
Copula-Granger [72]	Time series	$\mathcal{O}(n \times (\omega \times t^*)^2)$

network from time-lag features. A node is scored by breadth-first traversal on reversed 'following' edges. Visited neighbors' contribution is inverse-proportional to the geodesic distance. For the purposes of our simulation, we use sliding Euclidean distance alignment (e.g. analogous to cross-correlation) because LPD does not scale to the size of our simulations under DTW (see Table Table I). Finally, for influence maximization (IM), we use the independent cascade model for the 1-seed selection problem [36], on the network derived from [30]. The network describes the probability of any individual A sharing the same direction as B, and in the front of B.

For the Granger Causality method, we used the Copula-Granger approach that can be found in [72] to infer a causal network. Then, we convert the causal network to be a following network by designating X follows Y if the weight of Y Granger causes X is larger than the weight of X Granger causes Y.

To make leadership comparison possible, we report the global leadership rank ordered list for each method as follows. First, we create rank order lists for FLICA under PageRank. The FLOCK model, however, does not have the explicit ranking score, so we rank individuals based on decreasing time duration of leadership. Third, LPD assigns individuals with higher scores a higher rank. Finally, since IM uses the probabilistic network of influence, we construct the realization of this influence network. A node influences any node to which it has a directed path in the realized network. We rank individuals based on the expectation of nodes influenced by that node over 1000 realized networks. Lastly, for the Copula-Granger leadership framework, we use PageRank to evaluate the leadership ranking on the following network we created from the causal network.

5.3.4 Sensitivity analysis

Typically, real-world datasets are noisy. The high degree of noise can affect results of leadership inference. However, the effects of noise on leadership inference is unclear. Moreover, in our leadership framework, the main parameter is the time window ω. Using the wrong value of time window may affect the results as well. Hence, in this section, we consider the approach to measure the robustness of our framework.

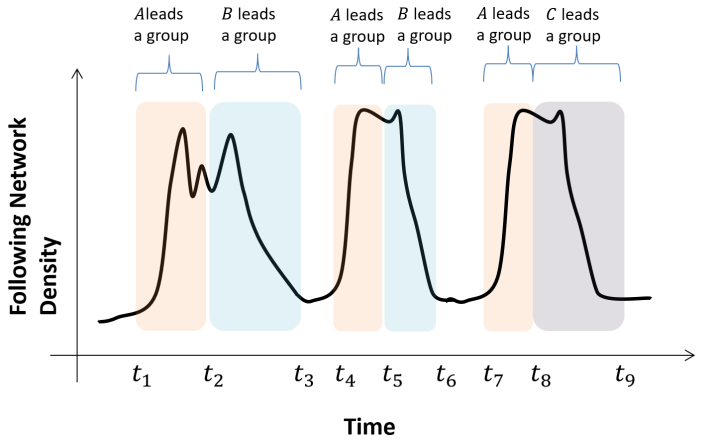


Figure 17. An example of time series of following-network density. There are three coordination events where individual A leads a group in first half of all coordination events while B leads a group after A for first and second coordination events. The third coordination event has C leads the second half of the event.

5.3.4.1 Support of faction leading

To measure the accuracy of initiator inference, we use a support value of being a leader of factions for each individual. A faction interval of L is defined as a sub-coordinated interval within a coordination event such that L leads a group. Given $\mathcal{F} = \{I_L\}$ is a set of faction intervals where I_L is a faction interval lead by L (The consecutive time interval that L leads the group) and a time window ω, a coordination event E_k is defined to be a combined interval of consecutive faction intervals. Specifically, if a faction interval I_i finishes at time t_i' while I_j starts at time $t_j \leq t_i' + \omega$, then both I_i and I_j factions are in the same coordination event.

Definition 18 (Coordination event) *Let $\mathcal{F} = \{I\}$ be a set of faction intervals and ω be a time window. A coordination event $E_k = [t_1, t_2]$ is a combined interval of consecutive faction intervals from $\mathcal{F}$. Any faction interval $I_i = [t_i, t_i']$ that occurs before another faction $I_j = [t_j, t_j']$ are in the same coordination event if $t_i' < t_j \leq t_i' + \omega$.*

Given $\mathcal{E} = \{E_k\}$ is a set of coordination events and $\mathcal{F}_L = \{I_i\}$ is a set of faction intervals lead by L, the support value of an individual L leading factions is defined as follows.

$$\sup(L) = \frac{|\{E|E \in \mathcal{E}, \exists I \in \mathcal{F}_L, I \subseteq E\}|}{|\mathcal{E}|} \tag{5.7}$$

The support value, $\sup(L)$, tells us the level of consistency that L happens to initiate its faction for any coordination event. If $\sup(L) \approx 1$, then it means L always initiates factions when coordination events occur. In contrast, if $\sup(L) \approx 0$, then there is a low chance that L initiates any faction.

Moreover, we can calculate a confident value of having A's and B's factions in the same coordination event given that A's faction occurs at the event as follows.

$$\text{Conf}(B|A) = \frac{|\{E|E \in \mathcal{E}, \exists I_i \in \mathcal{F}_A, \exists I_j \in \mathcal{F}_B, I_i, I_j \subseteq E\}|}{|\{E|E \in \mathcal{E}, \exists I \in \mathcal{F}_A, I \subseteq E\}|} \tag{5.8}$$

Figure 17 shows the toy example of coordination events in the form of density time series of following network. $\mathcal{E} = \{[t_1, t_3], [t_4, t_6], [t_7, t_9]\}$ is a set of coordination events. A set of faction intervals lead by A is $\mathcal{F}_A = \{[t_1, t_2], [t_4, t_5], [t_7, t_8]\}$, a set of faction intervals lead by B is $\mathcal{F}_B = \{(t_2, t_3], (t_5, t_6]\}$, and a set of faction intervals lead by C is $\mathcal{F}_C = \{(t_8, t_9]\}$. The

support values of A, B and C are $\sup(A) = 1$, $\sup(B) = 2/3$, and $\sup(C) = 1/3$ respectively. The confident values of having B's faction within a coordination event that has A's faction is $\text{Conf}(B|A) = 2/3$. But $\text{Conf}(A|B) = 1$ and $\text{Conf}(A|C) = 1$.

5.3.4.2 Simulation and accuracy measure

- **Initiator inference**: To measure the performance of framework vs. noise w.r.t. the initiator inference task, we use simulation datasets of movement time series within 2-dimensional space. A simulation dataset contains trajectories of individuals that have multiple coordination events (e.g. Figure 17). Each type of simulation dataset contains different type and level of noises. The task is to predict a support of each initiators as well as confident values. The performance of framework is determined based on the error between ground truth and predicted values of support and confident measures. If the framework performs well, the predicted support and confident value of initiators should be close to the ground truth.
- **Coordination interval inference**: To measure the performance of framework vs. noise w.r.t. the coordination interval inference task, we use simulation datasets of movement time series within 2-dimensional space. we compared the predicted faction intervals of each initiators to the ground truth directly. The result of comparison are reported in the form of true positive, false positive, and false negative values. True positive will be counted at time step t when a predicted and ground truth faction interval of initiator L occurs at time t. False positive will be counted when the framework predicts that a time step t is within L's faction but it is not. False negative will be counted when the framework predicts that a time step t is not within L's faction but it is.

5.3.4.3 Position and Direction noises

To measure the robustness of framework w.r.t. two tasks above, we consider two types of noises: position noise and direction noise. For a direction noise, instead of moving to a target direction at degree D compared to X-axis, an individual moves toward a direction $D + a$. The direction noise a is drawn randomly from a normal distribution with zero mean and γ standard deviation. For position noise, suppose (x, y) is the next position that an individual should move to, with position noise, the actual position that the individual moves is $(x + b_1, y + b_2)$. The position noise b_1, b_2 are drawn randomly from a normal distribution with zero mean and β standard deviation.

5.3.4.4 Time window sensitivity

A time window parameter ω is the main parameter of our leadership framework. we report the performance of framework when the time window is vary from the optimal time window. If the framework is robust, then it should perform well even when the time window value is set significantly different from the optimal value.

5.4 Results

5.4.1 Identifying leaders

In each simulation, we have the label of the true initiator(s). For each of the simulation trials, our method identifies the 'initiator' and 'rank ordered lists' (see Section 5.3.2). We set a window size ω by the TWIN heuristic [73] on the network density, window shift size $\delta = 0.1\omega$, and the λ threshold at the mean of the network density time series $\mathrm{d}(t)$.

Table II reports precision on PageRank rank ordered lists over all synthetic model simulations. We compare against previous works–FLOCK [30], IM [36], LPD [33], and Copula-Granger Causality Inference models [72]–which produce a single ranking over the entire trial.

The white rows in Table II report precision of leadership identification for a fixed initiator across all coordination events (global leadership). Gray rows report precision of initiator leadership where leaders change between pre-coordination and coordination intervals in the event (DM-S, HM-S), or precision of local leadership where the initiator changes per coordination event (EM). The rows labeled 'Top-4' report precision in identifying any of the multiple unordered initiators (CM, INIT-4) or precision *for the correct hierarchical order* (HM, HM-S).

On the white rows, FLICA is robust across all simulation models, while FLOCK, IM, and LPD perform well other than on INIT-4 simulations (e.g. with multiple initiators). However, in gray rows ("initiator switching") previous methods fail almost completely since they are unable to detect leadership prior to coordination. When the coordination state is more prevalent than the pre-coordination decision-point, ranking will favor an individual who happens to lead the dynamics in the coordination state (but may not have initiated the state). For Copula-Granger framework, it can infer correct initiators with high accuracy in Dictatorship model, while it fails for the most of models except INIT-4. This indicates that the Copula-Granger approach has a potential to infer leaders even though it is not designed to perform leadership inference.

The row reporting EM results is a special case of precision. Because we know each coordination event has a unique initiator, ranking individuals across all coordination events will

fail. Instead, we report precision in identifying the initiator of *each* coordination event. Since previous work generates only aggregate rankings, precision for these methods are not reported.

5.4.2 Case study: trained initiators in fish schools

We identify the top-k global initiators of the fish school trajectory dataset (see Section 4.2.2), where we have the labels of 'trained' individuals expected to lead the school to feeding sites. Table III reports precision of identifying trained fish as initiators over 24 trials. The Initiator column is precision of predicting a trained fish as a global initiator. The Top-4 rank column is precision of identifying trained fish as the top-4 ranking individuals. Similar to the simulation models, FLICA performs best overall, again suggesting that dynamic following network representation captures 'following' better than other features.

5.4.3 Case study: finding "initiators" of stock market events

We apply our leadership framework to stock market closing price data of the NASDAQ index (see Section 4.2.3). An 'initiator' in this context measures the extent that a stock increases or decreases in value before a large group of other stocks (e.g. a coordinated group). We apply the framework without any special consideration to the domain, only to qualitatively validate that we can discover known, large events.

Figure 18 shows the network density of the inferred 'following' network over time, where we discover coordination events with λ threshold at the 75th percentile of the network density time series. Pre-coordination and coordination intervals are shown in red and green, respectively. We find significant economic events such as the 2000 tech collapse, and 9/11. More interestingly, we discover known events which are reflected in the network density signal but not the NASDAQ

TABLE II

Precision of leadership identification on simulation models. (* indicates the $std \geq 0.1$).

Models/Methods	FLICA	FLOCK	IM	LPD	Copula-Granger
DM	**1**	**1**	**1**	**1**	**0.97***
HM (Top-4)	**1**	0.25	**1**	**1**	0.55*
LT	**0.99**	0.98*	0.99*	0.93*	0.07*
IC	**1**	**1**	**1**	0.99	0.45*
CM (Top-4)	**1**	**1**	**1**	0.99	0.69*
INIT-1	**1**	**1**	**1**	**1**	0.24*
INIT-4 (Top-4)	**0.74***	0.35*	0.51*	0.21*	**0.91***
DM-S	**1**	0	0.02*	0.25*	0.37*
HM-S (Top-4)	**1**	0	0.5	0.51	0.31*
EM	**0.92**	-	-	-	-
Random	0.01	0	0.01	0.17*	0.02*

TABLE III

Initiator identification precision in fish (* indicates the $std \geq 0.1$).

Ranking	Initiator	Top-4 rank
FLICA	**0.83***	**0.61***
FLOCK [30]	0.0	0.0
IM [36]	0.0	0.02
LPD [33]	0.17*	0.18*
Copula-Granger [72]	0.13*	0.10*

index. For example, we discover a technical econometric event, where the "TED Spread" (a surrogate of national credit risk) begins fluctuating in July 2007, and a small market failure in August 2011. Matching our intuition, the top-ranked companies in the coordination event associated with the year 2000 collapse are primarily in IT and semiconductors, including eBay and SanDisk in the top 10.

For the sanity check, we provided the results of the comparison between the following network density of NASDAQ stock market and a random walk one in Figure 19. We generated random-walk time series from the original NASDAQ closing price time series. Both original and random-walk versions shared the same distribution of difference between time steps, length, and the number of time series.

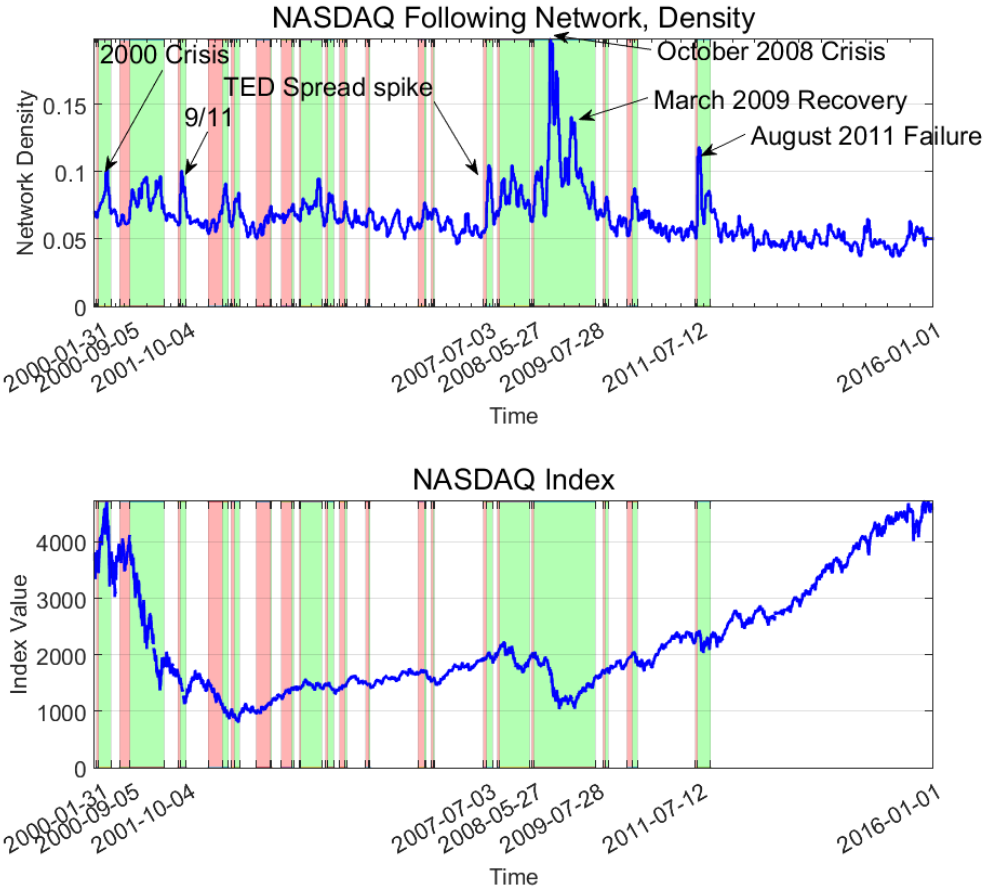


Figure 18. (Top) NASDAQ 'following' network density and (Bottom) NASDAQ index value. Pre-coordination and coordination intervals are shown in red and green, respectively. The framework detects many known events in financial data (labeled above). Many of these events are not reflected in the NASDAQ index.

For each time series of closing price X, we inferred the distribution $\mathcal{D}_X$ of the differences between the time steps. We created random-walk time series $\hat{X}$ that starts as the same price as

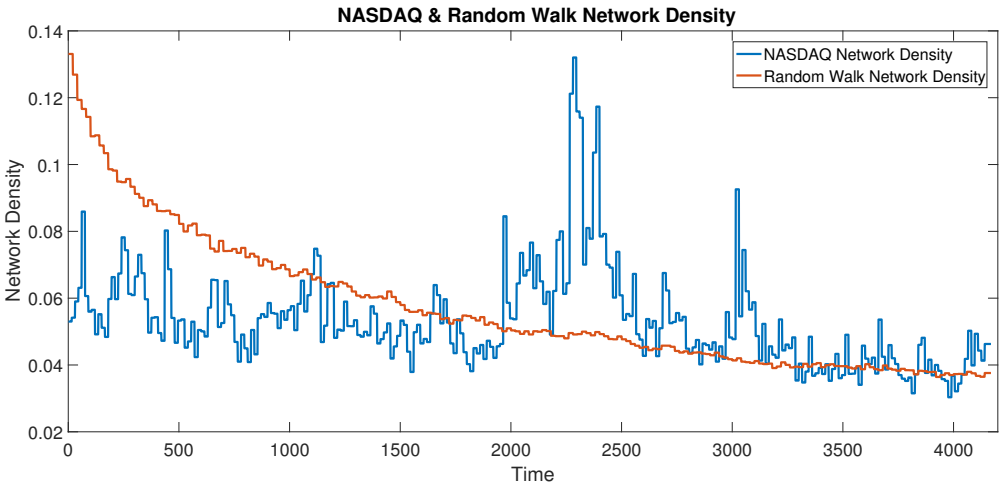


Figure 19. Comparison between time series of network density that generated from NASDAQ time series and from random walk time series.

the original time series, then we updated the next value of $\hat{X}$ by normally sampling a different value from $\mathcal{D}_X$. Hence, both X and $\hat{X}$ share the same $\mathcal{D}_X$.

We found that our following network density of NASDAQ is different from the random-walk one (Figure 19). Moreover, random-walk network density does not have any coordination events. This indicates that our network density can tell the difference between random-walk time series and the actual dataset that contains coordination events.

5.4.4 Case study: baboon activity classification by following network density

In this section, we demonstrate that our coordination events, which defined by following network density time series, corresponding to coordinated progression activity labeled by experts in baboon dataset (see Section 4.2.1).

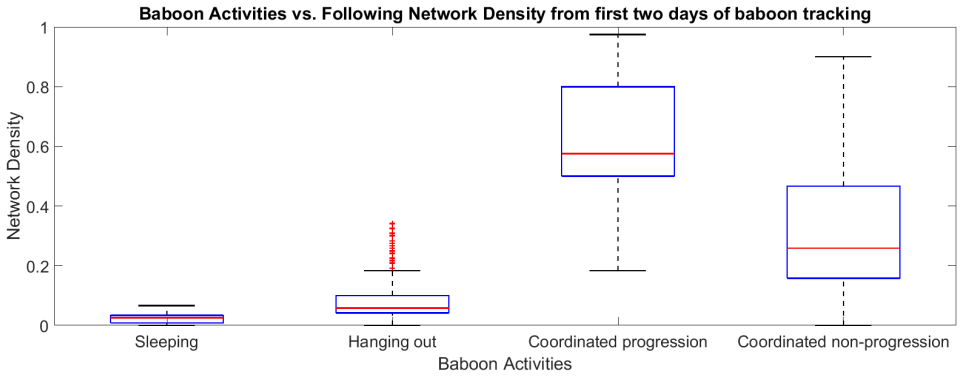


Figure 20. Baboon activities from first two days of tracking vs. following network density

For each time step, a group of baboon has an activity label as either sleeping, hanging out, coordinated progression, or coordinated non-progression. The group of baboons is considered to have a sleeping label when they are at their home tree to sleep. Hanging out activity happens when baboons stay around the same place without moving far away from the group. Coordinated progression is when a group of baboons having a strong coordinated movement to somewhere. Lastly, in coordinated non-progression, a group has a weaker coordinated movement than coordinated progression.

Figure 20 shows the distribution of network density for each activity. Sleeping and hanging out activities have low values of network density, which implies that the group rarely has following relations. On the contrary, coordinated progression and coordinated non-progression have high values of network density on average, which is similar to our coordination intervals.

This result illustrates that a following network density is informative with respect to the task of activity classification.

In the dataset, there are two days of baboon activity labels. We used the first day data to train a classifier to predict the second day activities as well as using the second day to train a classifier to predict the first day activities. We used Linear discriminant analysis (LDA) as our classifier and used only network density as a feature. We compared our result with Adversarial Sequence Tagging (AST) [67], with is the state-of-the-art method that performed activity labeling classification in the same dataset. Our aim is to show that a following network density is informative enough to make a simple classifier performs better than the state-of-the-art classifier method that used 24 features in the group activity classification task.

TABLE IV

Activities classification prediction accuracy in first two days of baboon data.

Method	Baboon Day 1	Baboon Day 2
Linear discriminant analysis (LDA) + following network density	**87.20%**	**70.82%**
Adversarial Sequence Tagging (AST) [67] + 24 features	77.30%	69.22%

The classification result is shown in Table IV. The accuracy results were calculated by the complement of hamming loss the same as [67]. By using only following network density as a feature, our simple classifier performs better than AST in both days.

5.4.5 Leadership model classification

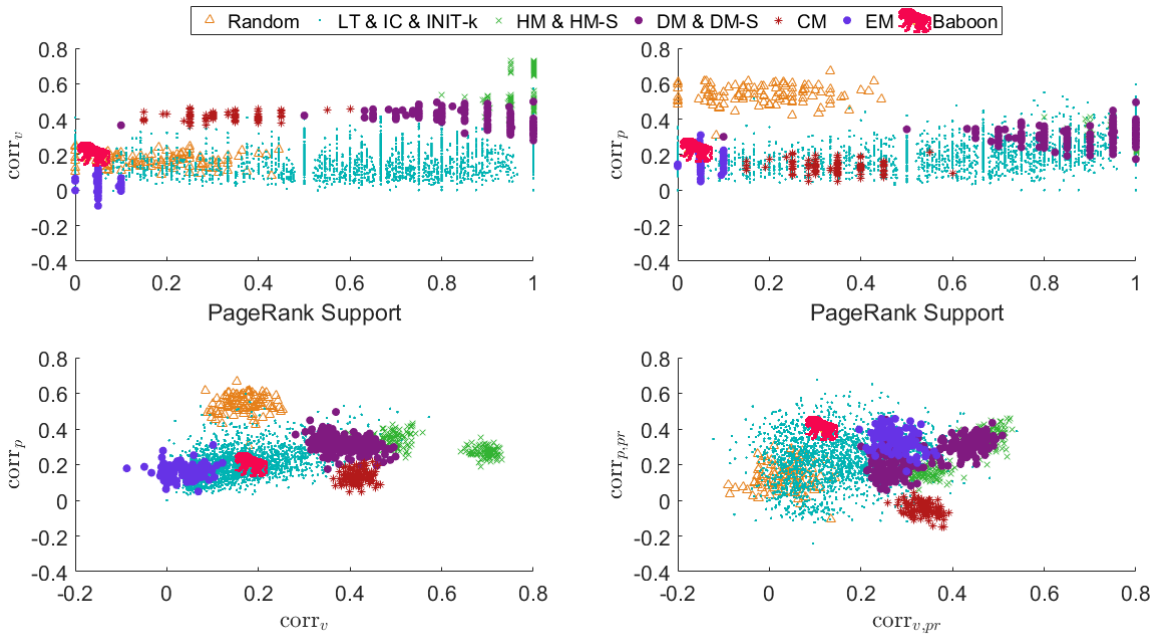


Figure 21. Comparison of feature spaces of leadership model classifications on simulations and real data

Recall, that we proposed several initiator rankings and ranking correlations (Section 5.2.5). Here, we do leadership model classification on each simulation trial using the proposed features derived from the rank correlations: corr_p, corr_v, $\text{corr}_{p,pr}$, $\text{corr}_{v,pr}$ and sup_{pr}. A classifier takes these features and produces a leadership model label, one per trial of the simulation model in the evaluation hold-out. We use 10-fold cross validation on Random Forests [74] over the 2700 total trials and report mean precision and recall across folds. Table V reports the classification results for each simulation model. We combine some models into a shared label because they share similar characteristics when we project them into our feature spaces. For example, DM, and DM-S models always have high corr_v but low corr_p.

Figure 21 visualizes sub-spaces of the full feature-space. Figure 21 (Top) shows the maximum support (sup_{pr}) over all individuals for this trial vs. the corr_v (the rank correlation between global and local VCH ranking) and corr_p. The sup_{pr} axis (x-axis) describes how 'dictatorial' (e.g. consistent) the leadership is across coordination events. DM therefore has high support, while EM (distinct leaders per coordination event) has low sup_{pr}. The corr_v and corr_p axes describe consistency between local and global convex hull rankings. HM has high velocity ranking because leaders accelerate in a consistent sequence, yielding consistent individuals movement outside of the VCH in the previous time step. The random model produces high corr_p because relative positions within the group are somewhat consistent. Therefore, a consistent set of individuals expand the PCH from the previous time step.

Figure 21 (Bottom-Right) reports the mean rank correlation between PageRank rank ordering, against PCH and VCH ranking *in each coordination event.* At the origin (0, 0), ranking

from the inferred 'following' network is uncorrelated with time series feature rankings in position or velocity. Following our intuition, the Random simulation has the lowest cross-domain feature correlation, while DM and HM have highest correlation between these domains. As the simplest simulations, DM and HM both dictate that leaders will have regular position (e.g. the front of the group), or velocity (accelerating in sequence before others). Simulations such as CM, LT, IC have indirect relationships between relative position and velocity vs. the following network ranking.

5.4.5.1 Baboon leadership model characterization

A key aspect of our simulation modeling is that we can characterize real datasets according to how they map into these feature-spaces, compared to synthetic models. We compute each rank correlations over high-confidence baboon events, labeled "Baboon" in Figure 21, thresholded at the 99th percentile of density. We observe that within different sub-spaces, the baboon ranking is similar to Random or Linear Threshold, and has low maximum support for global vs. local rank correlation features (e.g. corr_p). We see this rank correlation between both cross-domain axes (Figure 21 (Bottom-Right)). This suggests that in aggregate, baboon leadership is heterogeneous and context-driven, though overall closer to the Linear Threshold influence model (as biologically expected). This analysis provides a strategy for hypothesis testing and generation on contrasting time-scales and sub-spaces.

5.4.6 Following network density perturbation in baboon and fish data

In this section, we provided the results of the network density changes when individuals are removed from the datasets, either the high-rank ones or (uniformly) at random.

TABLE V

Random forest classification of synthetic leadership models using proposed features

Model	Precision	Recall	F1-score
DM, DM-S	0.86	0.80	0.81
HM, HM-S	0.69	0.98	0.80
LT, IC, INIT-k	0.99	0.97	0.98
CM	0.75	0.94	0.80
EM	1	0.54	0.64
Random	0.98	0.95	0.97

In the baboon dataset, we used leadership ranking during pre-coordination intervals to choose the top-k rank individuals. In the fish datasets, we used leadership ranking to choose top-k ranked individuals, which are also informed fish. For the randomly chosen individuals, we uniformly and randomly chose k individuals from the population. We repeated the random choice process 100 times and report the average of these results.

Table VI shows the average difference of the network density before and after removing k individuals from the baboon dataset. In the 'High-Rank' row, the network density decreases with the removal of high-ranked individuals. In contrast, in the 'Random' row, the network density is largely unaffected by the uniformly random removal of the individuals.

TABLE VI

The average difference of network density between before and after removing k individuals on the baboon dataset. We compare the case of removing high-rank individuals vs. random individuals. An element in the table represents the difference between the original network density and the network density after removing k individuals.

	Number of Individuals being removed from 16 individuals ($\#k$)			
	#1	#2	#3	#4
High-Rank	-5.45×10^{-03}	-7.37×10^{-03}	-9.65×10^{-03}	-10.14×10^{-03}
Random	0.057×10^{-03}	-0.004×10^{-03}	-0.504×10^{-03}	0.002×10^{-03}

In the fish data, Table VII shows the results of the network density difference after removing k individuals. In the 'High-Rank' row, the network density decreases with the removal of high-ranked individuals, albeit less so than in the baboon dataset. In the random case, the network density is still largely unaffected by the uniformly random removal of the individuals.

The results above show that by removing high-rank individuals in both baboon and fish datasets, the network density decreases significantly compared to randomly removed individuals. Moreover, we found that removing high-rank individuals in the baboon dataset resulted in a larger decrease of the network density than in the fish datasets. This suggests that baboons have a stronger following hierarchy than schools of fish.

TABLE VII

The average difference of the network density between before and after removing k individuals from the fish data. We compare the case of removing high-rank individuals vs. random individuals. An element in the table represents the difference between the original network density and the network density after removing k individuals.

	Number of Individuals being removed from 70 individuals ($\#k$)			
	$\#2$	$\#4$	$\#6$	$\#8$
High-Rank	-20.26×10^{-04}	-3.96×10^{-04}	-2.45×10^{-04}	-8.44×10^{-04}
Random	-0.55×10^{-04}	-0.04×10^{-04}	-3.90×10^{-04}	-3.88×10^{-04}

5.4.7 Sensitivity analysis

We conducted the sensitivity analysis based on Section 5.3.4 to demonstrate the robustness of our framework. We used simulation datasets that have the time delay for following relations less than 30 time steps. Hence, the optimal time window ω is 30 time steps. In the initiator inference, the results of loss values of the initiator-support prediction are shown in Figure 22. Each cell in these three sub figures represents a loss value which is the difference between the ground-truth and the predicted value of the initiator support (Eq. Equation 5.7). In general, unsurprisingly, the loss value increases with noise.

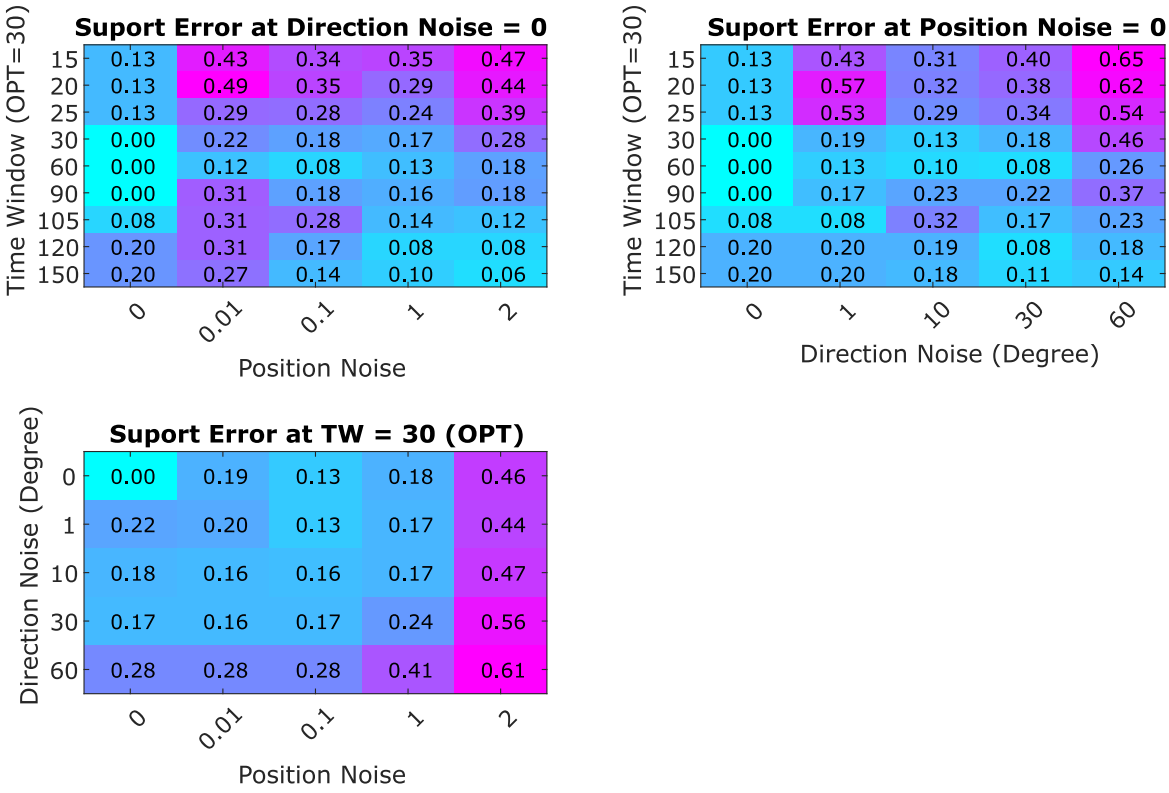


Figure 22. Loss values of prediction of the initiator support for different levels of noise and time window sizes. A lower value implies a better prediction result.

Figure 22 (below) illustrates that both the position and direction noise affect the prediction performance. In top-left and top-right plots, they show that when we set the time window below the optimal value ($\omega < 30$), the loss of support inference is significantly higher than setting the time window above the optimal time window. This suggest us to try to guess the possibly

maximum time delay in datasets and avoid setting the time window parameter below this value if the ground truth regarding time delay is not available.

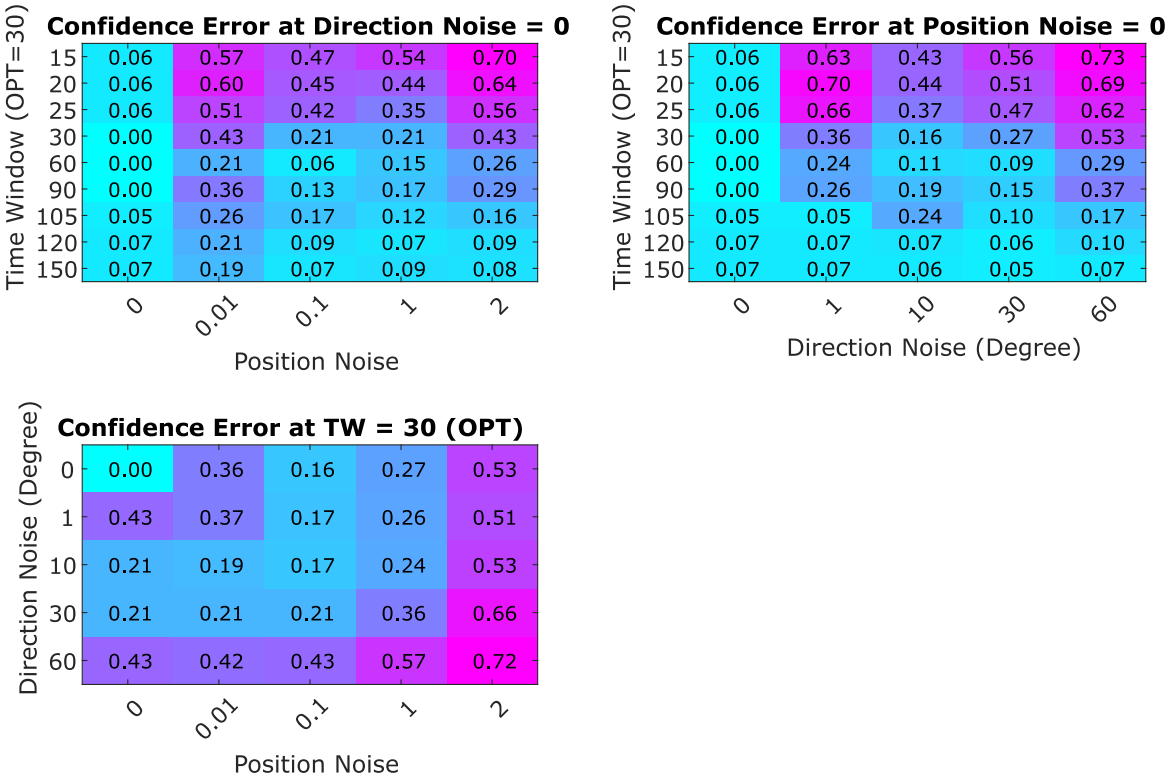


Figure 23. Loss values of prediction of the initiator confidence for different levels of noise and time window sizes. A lower value implies a better prediction result.

Figure 23 shows the loss values of the initiator-confidence prediction. Each value in these three sub figures represents a loss value, which is a difference between the ground-truth and the predicted confidence values. Similar to the support result, higher levels of noise results in higher loss values. Similarly, setting the time window below the optimal value also severely affects the framework performance in confidence prediction.

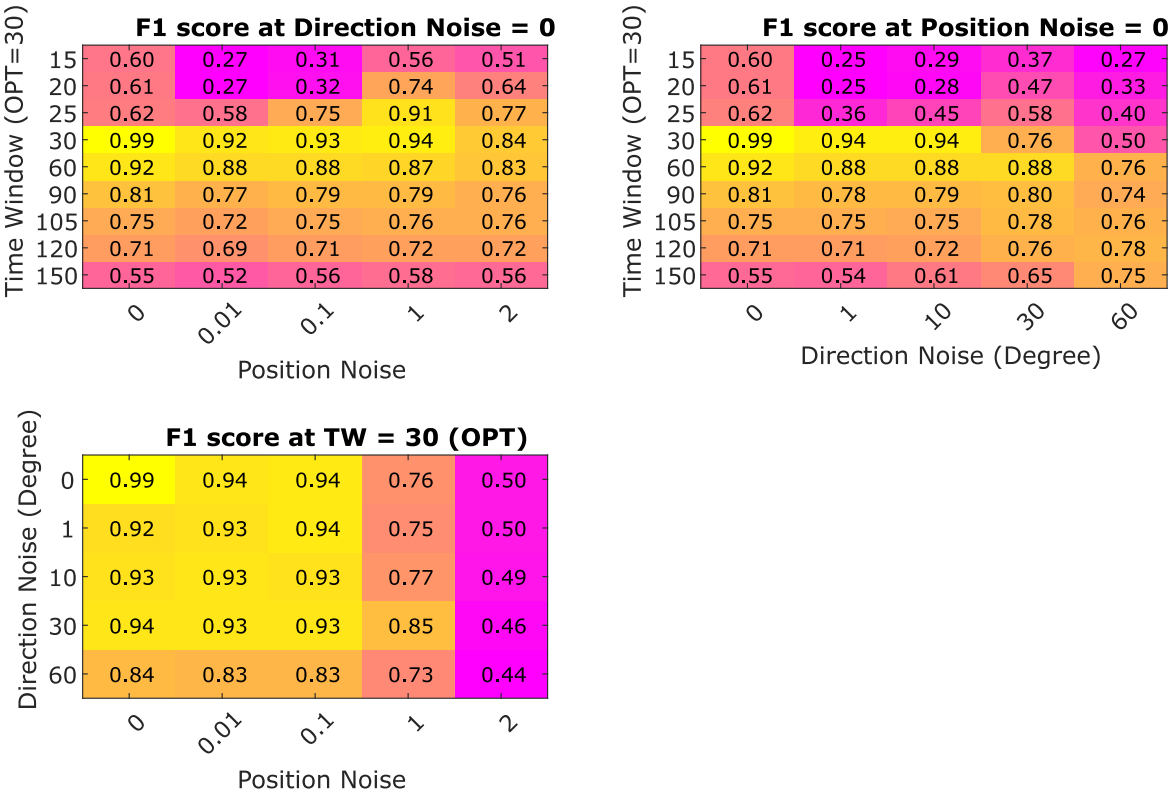


Figure 24. F1 scores of coordination inference for different levels of noise and time window sizes. A higher value implies a better prediction result.

In the coordination inference task, Figure 24 shows the F1 scores of the coordination inference for different levels of noise and time window sizes. Each value in these three sub figures represents the F1 score of the prediction. Similar to the previous results in the initiator inference, higher levels of noise reduce prediction performance. Setting the time window below the optimal value also decreases the prediction performance severely.

In conclusion, when we increase the amount of noise in the datasets, the framework performance decreases. Moreover, we found that if we set the time window parameter below the optimal, it severely affects the framework performance in both initiator and coordination inference tasks. In contrast, when we set the time window above the optimal value, the framework performance drops only slightly. The optimal time window is, then, a value above the possible highest time delay that two individuals can follow each other in the dataset. This suggests that the users of our framework should try to come with an educated guess of the possible values of time delays that are relevant for the application context and set the time window parameter accordingly.

5.5 Discussion

We narrow the gap between the biosociological view of leadership in group decision-making and the computational approaches to leadership inference. The work presented in this paper formalizes a **new computational problem**, namely Coordination Initiator Inference Problem, and proposes the concrete, simple yet powerful, unsupervised general framework as a solution. The framework is capable of (1) identifying events of coordinated group behavior, (2) identifying leaders as initiators of these events, and (3) classifying the type of leadership

process at play. We validate the accuracy of our framework in performing all three of these tasks using 2,700 simulated datasets. Since there are no methods for local leadership inference and leadership model classification, we compared our framework with the state-of-art methods for global leadership identification. Our method performance is consistently competitive and its abilities go beyond other approaches in all datasets. We further show that the framework can provide insights on real-world data, including data on collective animal movement and the economy. The methodology presented here is general and applicable to a wide variety of domains where coordination across many individuals or entities is observed. Moreover, our framework is highly flexible, and can easily be extended to incorporate other models of leadership or other features used in model classification, depending on the details of the system being analyzed. For reproducibility, we provide our code and simulation datasets at [75]. In the next chapter, we will extend the FLICA framework to cover the case of multiple coordinated groups that occur concurrently. The new proposed framework can be used to infer these concurrent groups from time series.

CHAPTER 6

MFLICA: A FRAMEWORK FOR MULTIPLE-FACTIONS LEADER IDENTIFICATION IN COORDINATED ACTIVITY

6.1 Introduction

FACTION INITIATOR INFERENCE PROBLEM: To reach collective goals, group's members must coordinate with each other. Multiple factions within a big group may exist solving their sub-tasks in helping the entire group achieve the collective goals. **Given time series of individual activities, our goal is to identify periods of coordination and the subsequent coordinated activity, find factions of coordination if more than one exist, as well as identify leaders of each faction**

In this chapter, we present a framework for multiple-factions leader identification in coordinated activity (mFLICA), which is the extention framework of FLICA from the previous chapter. The steps of dynamic following network inference are the same as FLICA. In addition to FLICA, mFLICA has two more steps to infer factions and an appropriate time window.

This chapter has been previously published in Chainarong Amornbunchornvej and Tanya Y. Berger-Wolf. 2018. Framework for Inferring Leadership Dynamics of Complex Movement from Time Series. In Proceedings of the 2018 SIAM International Conference on Data Mining (SDM), pp. 549-557. https://doi.org/10.1137/1.9781611975321.62 See Appendix A for the copyright permission document from the publisher.

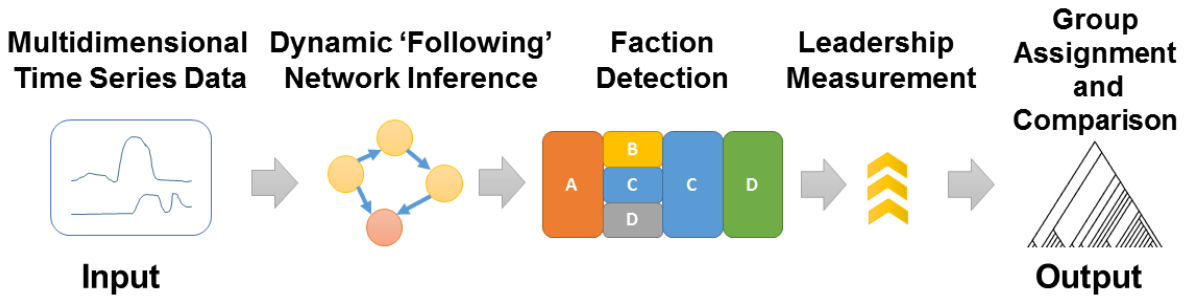


Figure 25. A high-level overview of mFLICA framework

We propose the following framework to solve Faction Initiator Inference Problem. The framework is designed to infer a set of factions, faction intervals, and their initiators from time series. Figure 25 depicts the overview of our framework.

In our setting, we have a set of n m-dimensional time series $\mathcal{U}$ and a time window ω as inputs of our framework. The time series $U \in \mathcal{U}$ can be in any domain (e.g. position, stock market closing price, or actions in social media). The size of $\mathcal{U}$ is $[n \times m \times t^*]$ where t^* is a length of time series within $\mathcal{U}$. Then, our framework proceeds in the following steps. The first step is to measure following relations (see Section 5.2.1) among time series. Then, the framework constructs a dynamic following network based on these following relations (see Section 5.2.2).

Afterward, the framework uses the dynamic following network to detect factions and their initiators based on the network structure (see Section 6.2.1). For each faction, we rank time series based on the network structure of faction's members by PageRank (see Section 6.2.3). Finally, we report time series of faction assignment as well as faction's rank ordered lists.

Moreover, without prior knowledge of a time window ω, we propose a method to infer an appropriate time window ω^* from $\mathcal{U}$ in Section 6.2.2.

6.2 Methods

6.2.1 Factions detection and coordination intervals.

For each following network $G = (V, E)$, factions are network components such that all member nodes directly connect to their initiator (Lemma 3.3.1). We infer factions based on Definition 10 and the coordination intervals of factions are discovered based on Definition 11.

According to Lemma 3.3.1, initiator nodes have outgoing-degree zero, and all nodes within the similar faction directly connect to their initiator. However, due to the introduction of the time window ω, some nodes might not have direct edges to the initiators. Therefore, we relax the constraint of faction membership to make all nodes which have any directed path to an initiator to be members of the initiator's faction.

Since a faction is a directed connected component where all nodes are reachable from the initiator by inverse paths, we use Breadth-First Search (BFS) to identify all reachable nodes from each initiator node in the following network in order to find members of each faction. The pseudo code of this step is in the supplementary material.

A useful statistic about factions (used later) is the *faction size ratio*. Let $G_l = (F_l, E_l)$ be an induced subgraph of G defined by faction F_l, then the faction size ratio of F_l is defined as follows:

$$\text{fs}(F_l) = \frac{|E_l|}{\binom{|V|}{2}}. \tag{6.1}$$

6.2.2 Time window inference.

In reality, some following relations might not be cause by explicit initiators since they either happen by chance or are due to other factors which are not related to the influence of leaders. For instance, if a follower is unable to observe a leader's pattern, then the follower cannot be influenced by the leader. Different types of time series have different limitation of 'observation memory', which is the limitation of time delay Δ such that a follower can truly observe and imitate its leader's actions or can get commands from a leader.

Hence, to represent the concept of observation memory limitation, we set the time window ω to limit the length of the time delay Δ that can measure following relations. Moreover, ω helps us prevent the comparison of time series between different coordination events.

Nevertheless, if we set ω too small, we miss inferring some following relations that have $\Delta > \omega$. On the contrary, long-length ω causes false positive matching between repeated patterns of different coordination intervals. Therefore, a proper ω^* should be able to infer a higher number of true following relations than any ω. Even if some random following relations might appear when we choose ω instead of ω^*, this is not an issue. Since these random following relations appear by chance and with lower probability, they have a relatively small effect on the number of following relations.

In our framework, without the knowledge of ω, we use ω that maximizes the average coordination measure Ψ (Equation 3.2). Given a dynamic following network based on the time window ω, for each time step t, we calculate Ψ_t by designating each faction to be a cluster and creating the last cluster for all time series, which are not in any faction. Then, $\hat{\Psi}_\omega$ is computed

from the median of $\{\Psi_1, ..., \Psi_{t^*}\}$. $\hat{\Psi}_\omega$ is used to be a representative coordination measure value of ω. Hence, the optimal ω^* is computed as follows:

$$\omega^* = \underset{\omega}{\text{argmax}}(\hat{\Psi}_\omega). \tag{6.2}$$

6.2.3 Leadership comparison.

There are several methods that are widely used for ranking important nodes within the graph. One of the well-known methods that consider the higher-order relation within a graph is PageRank [66]. In our approach, we deploy PageRank (Equation 3.1) on the following network to rank individuals within each faction and report the rank ordered lists for each time step. Even though PageRank scores are computed from the entire network, we compare individuals' ranking score only within the same faction and create a rank order list for each faction.

6.3 Evaluation Datasets

6.3.1 Leadership models.

The evaluation of the framework is conducted based on four models of coordination mechanisms: Dictatorship model (DM) (Section 4.1.1), Hierarchical model (HM) (Section 4.1.2), Independent Cascade model (IC) (Section 4.1.7), and Crowd model (CM) (Section 4.1.5).

6.3.2 Synthetic trajectory simulation.

We generate time series datasets based on the models described above. For each dataset, it consists of 30 individuals' time series of X, Y coordinates. Each time series has a length of 4,000 time steps. A coordination event consists of multiple faction intervals, described below.

We have five coordination events for each dataset. For each model above, the coordination event can be divided into two types.

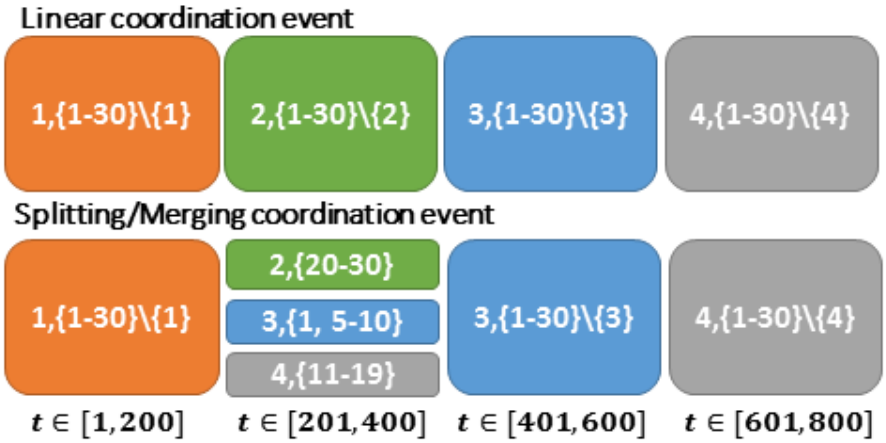


Figure 26. Linear (above) and Splitting/Merging (below) coordination event. Each block represents a faction such that the first element is the leader ID and the second element is the member IDs set. The time interval each faction appears is at the last line.

6.3.2.1 Linear coordination event.

There are four factions for each coordination event. The first faction has ID(1) as a leader and others are followers. This faction lasts for 200 time-steps. The next faction is lead by ID(2) and its coordination interval is $[201, 400]$. The third faction appears within $[401, 600]$ interval and it has ID(3) as a leader. In the last faction, ID(4) leads the group to stop moving and

the group completely stops moving around time step $t = 700$. Everyone stops moving within $[700, 800]$, then the group proceeds to the next coordination event again.

6.3.2.2 Splitting/Merging coordination event.

In this type of coordination event, splitting and merging of factions happens. Within the $[1, 200]$ interval, ID(1) leads a single faction with its direction vector. Then, at $t = 201$, the group is split into three factions and they appear within $[201, 400]$ interval. The first faction is lead by ID(2) and about a third of the previous faction members are followers (Figure 26 below). The ID(2) has its own direction vector. ID(3) leads the second faction with another one third members from the previous faction. ID(3) has a different direction from ID(2). Lastly, ID(4) leads the rest of the individuals. ID(4) also has its own direction, which is different from ID(2)'s and ID(3)'s.

At $t = 401$, the factions lead by ID(2) and ID(4) are merged into the faction of ID(3); ID(2) and ID(4) follow the ID(3)'s direction. At the $[401, 600]$ interval, ID(3) leads all the individuals. Finally, ID(4) leads the faction to stop moving between $t = 601$ and $t = 700$. The group completely stops at the $[701, 800]$ interval. Note that leaders in each faction are informed individuals in the Crowd Model. Instead of having only one leader for each faction, we have three informed individuals in the Crowd Model.

For each leadership model and its coordination event type, we generated 100 datasets. In total, each model has 200 datasets except IC, for which we explore all nine possible combinations of parameters. In total, we have 1,800 datasets for the IC model.

6.3.3 Biological datasets

6.3.3.1 Baboon trajectories.

The data consists of time series of latitude-longitude location pairs for each baboon every second. 16 individuals whose collars remained functional throughout the time are analyzed for a case study of a merging coordination event. See Section 4.2.1 for more information.

6.3.3.2 Fish schools trajectories.

The fish dataset is a set of time series of fish positions from a video record of a school of golden shiners (*Notemigonus crysoleucas*). See Section 4.2.2 for more information.

6.4 Evaluation criteria

For each simulation dataset, we have the ground truth of an individual's membership in a faction and the identity of the faction's leader. We compared the inference result from each method against the known ground truth to evaluate the method's performance.

6.4.1 Individual assignment.

For all models, for each time step, the accuracy of the individual assignment is the number of inferred individuals' factions that agree with the ground truth, divided by the total number of individuals. Note that, in the Crowd Model, each faction F has a set of informed individuals and individuals belong to F if they follow any informed individual in F.

6.4.2 Leadership prediction.

For all models except the Crowd Model, the true positive TP is the number of inferred leaders who are indeed the ground truth leaders. The false positive FP is the number of

inferred leaders who are not the actual leaders. The false negative FN is the number of actual leaders who are inferred to be non-leaders. In the Crowd Model, TP is the number of inferred leaders who are informed individuals from the right faction. FP is the number of leaders who are uninformed individuals. FN is the number of ground truth factions such that all informed members are non-leaders. We calculated F1-Score to estimate the performance of the leadership prediction for each framework.

6.4.3 mFLICA Time complexity

Let n be a number of time series, ω be a time window, δ be a shifting factor (we use $\delta = 0.1\omega$), and t^* be a total length of time series. By deploying DTW Sakoe Chiba band technique [76] setting δ as a band limitation, the time complexity of computing a following network is $\mathcal{O}(n^2 \times \omega \times \delta)$. Since we need warping paths, not a distance, the upper/lower bounds tricks which are used to speed up DTW found in the time series literature cannot be applied here. The number of following networks we need to compute is $\frac{t^*}{\delta}$. In total, the time complexity of our framework is $\mathcal{O}(n^2 \times \omega \times t^*)$. Additionally, we might explore k candidates of ω in order to find the optimal ω. Since k is a constant, the asymptotic time complexity of our framework also remains the same. This expensive cost is unavoidable and it makes our framework hard to be a scalable framework.

6.5 Results

6.5.1 Leadership Identification.

For each simulation model in Section 6.3.1, we evaluated results from all datasets using the criteria in Section 9.4. We set ω time window by the method from Section 6.2.2 and set time

TABLE VIII

Factions and Leaders identification on simulation models

	Leadership F1-score		Assignment Acc.	
Dataset	mFLICA	FLOCK	mFLICA	FLOCK
DM-L	**0.94**	0.92	**0.89**	0.86
DM-MS	**0.94**	0.91	**0.86**	0.84
HM-L	**0.94**	0.91	**0.94**	0.86
HM-MS	**0.95**	0.90	**0.86**	0.81
IC-L	**0.91**	0.86	**0.86**	0.80
IC-MS	**0.89**	0.85	**0.79**	**0.79**
CM-L	**0.82**	0.64	**0.83**	0.64
CM-MS	**0.75**	0.67	**0.64**	0.55

TABLE IX

Rank orders median accuracy within factions

	Top3 Rank Order Accuracy	
Dataset	mFLICA	FLOCK
HM-L	0.75	0.78
HM-MS	0.72	0.76

shift $\delta = 0.1\omega$. The results of faction assignments and leaders identification are in Table VIII. Each row with the label '-L' is a model with Linear coordination event type (Section 6.3.2.1) and '-MS' represents a model with Splitting/Merging coordination event type (Section 6.3.2.2). The 2nd and 3rd columns represent the results of leadership prediction F1-Scores of mFLICA (our proposed framework) and the modified FLOCK framework [30, 77], and the values in these columns are calculated from the median of all datasets from a given leadership model. The 4nd and 5rd columns represent individual assignment accuracy results. We took the median of all given-model datasets to represent each model accuracy. Unsurprisingly, mFLICA beat FLOCK in all models. The result implies that the simple framework like FLOCK has a limitation when it needs to deal with complicated noisy leadership models.

TABLE X

A school of fish inference median accuracy over 24 trails

Method	Trained fish factions	Trained fish leaders
mFLICA	0.90	0.88
FLOCK	0.37	0.27

In hierarchical models, we reported the result of top 3 rank order inference accuracy within each faction in Table IX. The table rows represent leadership model datasets. The columns are accuracy, which determined by the percentage of top-3 individuals from the ground truth appear in the list of top-3 inferred list. Even though mFLICA has a competitive results, the FLOCK framework performs better, which makes sense since the hierarchical model has a linear hierarchy structure and the leader is always in the front of the group's direction, which matches the fundamental assumption of FLOCK.

6.5.2 Case study: trained leaders in fish schools.

We considered any fish within the faction of a trained fish to be following the trained fish. Among 24 trails of fish movement, the medians of inference accuracy of a fish following the trained fish are in column 2 in Table X. We also measured the accuracy of inferred initiators being the trained fish in each trial (column 3 in Table X). According to the results in Table X,

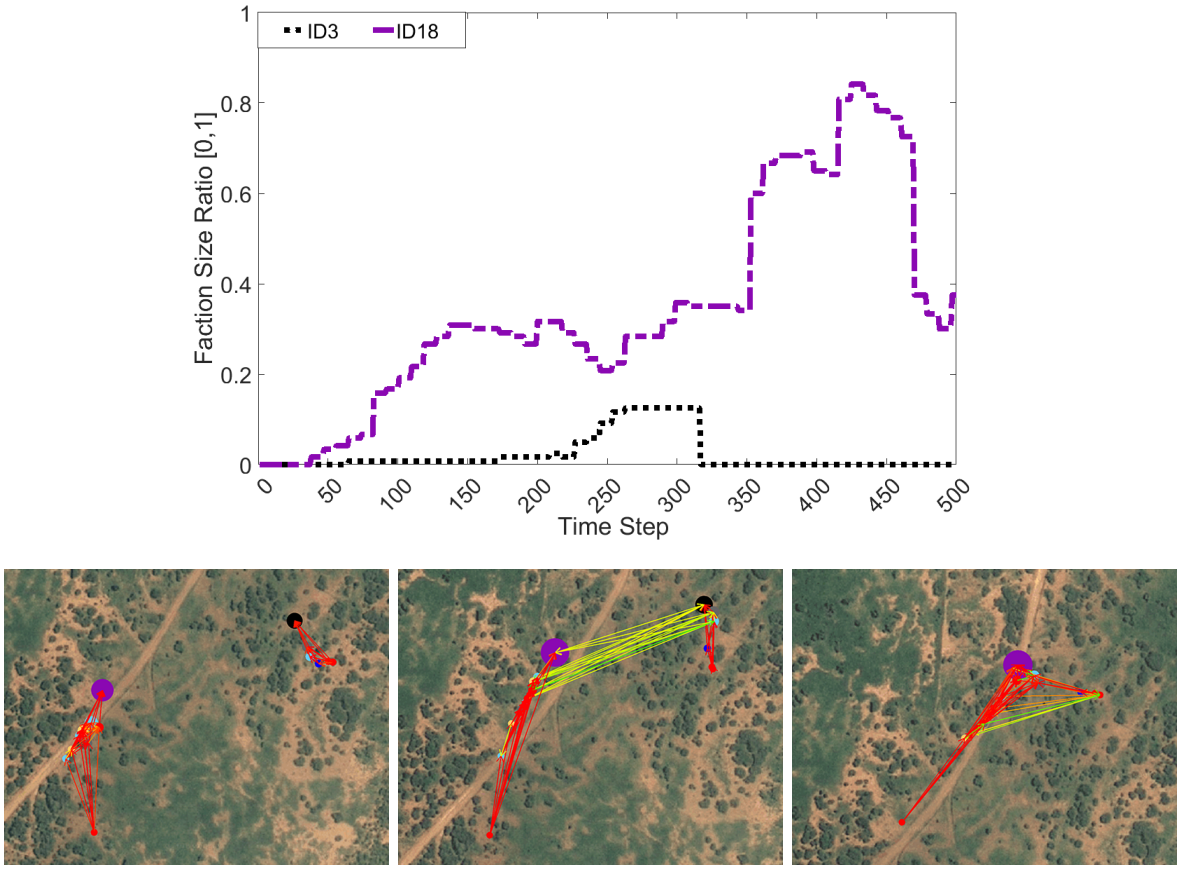
mFLICA performs significantly better than FLOCK in both aspects. This is because the fish datasets are tremendously noisy, and the DTW in mFLICA is more robust to the noise than the simple FLOCK model [33].

6.5.3 Case study: detecting the group merging event of baboons.

We used a baboon dataset (see Section 4.2.1) to demonstrate an example application of our framework to find transitions of coordinated events in real datasets. We focused on the dataset during the period when the merging of two groups happens on Aug 3, 2012, 08:49:01 AM. The length of the trajectories is 500 seconds. Figure 27 illustrates the result when the merging happens. Before time $t = 300$, a faction lead by ID(3) (black node) starts moving in the same direction as the faction lead by ID(18) (purple node). The process is measured by the Faction size ratios (Equation 6.1) of both factions, which increase over time. After $t = 300$, ID(3) faction is merging with ID(18)'s faction to become a single faction at $t = 400$. After merging, because the faction of ID(18) gains more members, its Faction size ratio (Equation 6.1) increases. Hence, by observing Faction size ratios lead by each individual, we can find merging events (or spiting events).

6.5.4 Centrality measures in multi-faction datasets

In this section, we explore the use of centrality measures to infer faction initiators. We used 200 simulated datasets from the dictatorship model to conduct the analysis. For each dataset, we created a global static following network and used centrality measures on this network. In each dataset, we have 30 individuals and four of them are initiators.



(b) $t = 300$ (c) $t = 350$ (d) $t = 400$

Figure 27. The merging coordination event. (Top) Faction size ratios (Equation 6.1) of ID(3) and ID(18) factions. (Bottom) The GPS locations of individuals in the map over three different time steps ($t = 300, 350, 400$), with the 'following' network, and PageRank indicated by node size. ID(3) is black and ID(18) is purple. The red edges have higher edge weights than the light edges.

TABLE XI

Jaccard similarity between top-4 ranking individuals from centrality measures and the ground truth set of four initiators in dictatorship model from 200 datasets.

	Centrality methods		
Event types	PageRank	IN-Degree	Closeness
Linear	0.85	0.84	0.54
Merge/split	0.64	0.67	0.53

The Jaccard similarity result between the top-4 ranking individuals from the centrality measures and the ground truth set of four initiators is in Table XI. PageRank and IN-Degree centrality perform well in dataset containing the simple linear coordination events while closeness centrality performs the worst. This is because initiators in this setting are supposed to have a higher number of followers than non-initiator individuals, which implies the higher ranking w.r.t. PageRank and In-Degree centrality. In contrast, initiators are not necessary close to their followers in the network, which made closeness centrality perform poorly. For the datasets that contain merge/split-coordination events, since there is a complicated dynamics of interactions among the factions, the simple centrality measures fail to capture the true initiators altogether.

Table XII illustrates the result of supports of four initiators being in the list of individuals ranked top-4 by the centrality measures in linear-coordination datasets. Similarly, PageRank

TABLE XII

Support of four initiators being in the list of top-4 ranking individuals from centrality measures in 100 datasets containing linear coordination events.

	Centrality methods		
Initiator's ID	PageRank	IN-Degree	Closeness
ID1	1	1	0.92
ID2	1	1	0.66
ID3	1	1	0.36
ID4	0.39	0.37	0.20

and In-Degree centrality perform well, while closeness centrality perform poorly. For the merge/split-coordination datasets, Table XIII shows that all centrality measures perform poorly to infer ID2 and ID4 initiators while they perform well to include ID1 and ID3 in their top-4 ranking lists. This is because ID1 and ID3 spent significantly more time leading their factions than ID2 and ID4.

In conclusion, these results emphasize the need of a dynamic following network approach to deal with the complicated problem of inferring the initiator of a faction.

6.6 Discussion

In mFLICA work, we formalized the Faction Initiator Inference Problem and provided an end-to-end general, unsupervised framework as the novel solution that can be used to study a wide range of coordinated activities. The framework is competitive against a non-trivial baseline method in both simulated and real-world datasets. Moreover, we demonstrated that the framework can be used to identify merging events as well as factions and initiators at each time step in biological datasets. This example implies that our framework opens opportunities for scientists to ask questions about coordinated activities and is able to create scientific hypotheses and test them. Our framework is powerful and almost parameter free (we need only

TABLE XIII

Support of four initiators being in the list of top-4 ranking individuals from centrality measures in 100 datasets containing merging/splitting coordination events.

	Centrality methods		
Initiator's ID	PageRank	IN-Degree	Closeness
ID1	1	1	1
ID2	0.29	0.47	0.20
ID3	1	1	0.83
ID4	0.28	0.19	0.08

a similarity threshold σ and time shift δ parameter). The scalability bottleneck is the DTW method used to compare time series. The existing DTW lower/upper bound techniques cannot be applied directly in our case since they only compute the distance between time series and not the actual wrapping path needed in our framework. With simpler and faster similarity computation, our framework can become highly computationally scalable. In the future, such more scalable approaches should be investigated. Another future work we plan to explore is the causality inference, which is closely related to leadership inference in the sense that initiators cause their followers' actions. We are planning to report the Granger causality results for leadership inference in our next paper. The code, datasets, and supplementary files that we used in this paper can be found at [78]. In the next chapter, we will extend the FLICA framework to make it to be able to infer traits of leaders from the concept of convex hulls in movement features.

CHAPTER 7

TRAITS OF LEADERS FRAMEWORK

7.1 The Proposed Approach

As stated earlier, there are many aspects of leader identity that may be used as the defining traits of the leader: the individual may be the oldest, biggest, wisest, or loudest. However, here we focus on the *behavioral* aspects of successful leadership, particularly in movement initiation. Are the leaders the ones who move first, move in the new direction, stay at the front, etc.? These are aspects of leadership behavior that also inferrable from the spatio-temporal time series data directly. We use the notion of a convex hull of the variable of interest for the the group versus an individual, particularly the leader individual.

7.1.1 Bidirectional agreement in multi-agent systems

The use of convex hull to analyze traits of a leader in this paper is motivated by the work on bidirectional agreement dynamics in multi-agent systems by Chazelle [47]. Chazelle showed that in a multi-agent system, the states of all individuals converge to a group consensus if each individual changes its state for each time step under what he calls the "Bidirectional Agreement

---

This chapter has been previously published in Chainarong Amornbunchornvej, Margaret C.Crofoot, and Tanya Y. Berger-Wolf. 2017. Identifying Traits of Leaders in Movement Initiation. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (ASONAM), pp. 660-667. https://doi.org/10.1145/3110025.3110088. See Appendix A for the copyright permission document from the publisher.

condition". The Bidirectional Agreement condition constrains an individual's choice of the state at each time step within the convex hull of the states of its neighbors (in the arbitrary agent network) in the previous time step. Thus, to break the group consensus state some individual must break the convex hull condition at some time point. In the collective movement context, a state of an individual at time t can be the individual's position, direction of movement, velocity, or acceleration.

Initially, all individuals' state are within the convex hull of the group's state. However, when the group initiates movement, the group changes its state from the initial state to unstable state. Leaders who initiate movement must break the convex hull of the group state to change the state of the group from one state to another state. After movement initiation, under "Bidirectional Agreement condition", the group converges to a stable state and everyone stays within the convex hull of the group's state again unless the convex hull is breached. In other words, we hypothesize that leaders are the state changers who start breaking the convex hull before others, and we test that hypothesis. For example, suppose we define a state as a position of each individual. Initially, by definition, all individuals are within the convex hull of an individual's positions. When leaders initiate movement by leading at the front, they must step outside the convex hull of group's positions.

We can define individual states to be any variable directly derivable from the time series data, including individual positions, velocities, or directions. However, the question is which of these variables' convex hull of the group state that leaders actually break when they initiate

movement that everyone follows. In this paper, we consider the states that leaders changes as behavior traits of leaders and we aim to infer these traits from time series data of group movement.

7.1.2 Bidirectional Agreement condition

First, we start with the one-dimensional states of Bidirectional Agreement condition. At any time t, suppose $S_t = \{s_1(t), \ldots, s_n(t)\}$ is a set of individual states at current time where $s_i(t) \in \mathbb{R}$ is a one-dimensional state of individual i at time t, $m_i(t)$ is a point in S_t that is closest to $s_i(t)$ but has a value at most $s_i(t)$, $M_i(t)$ is a point in S_t that is closest to $s_i(t)$ but has a value at least $s_i(t)$ and a constant $\rho \in (0, 1/2]$. The work by Chazelle [47] defines a Bidirectional Agreement condition as follows:

$$(1-\rho)m_i(t) + \rho M_i(t) \leq s_i(t+1) \leq \rho m_i(t) + (1-\rho)M_i(t). \tag{7.1}$$

The interpretation is that if all individuals change their state within the bound of their neighbor's states, the group will converge to a collective state, which is a stable state. In high-dimensional states, the Bidirectional Agreement condition still requires individuals to change their state from time to time within the bound of their neighbor's states, which is a convex hull of neighbor's states to make the group converge to a stable state. In other words, if all individual states always stay within their neighbor-state convex hull, then the group converges to a single point of collective state and stay there forever.

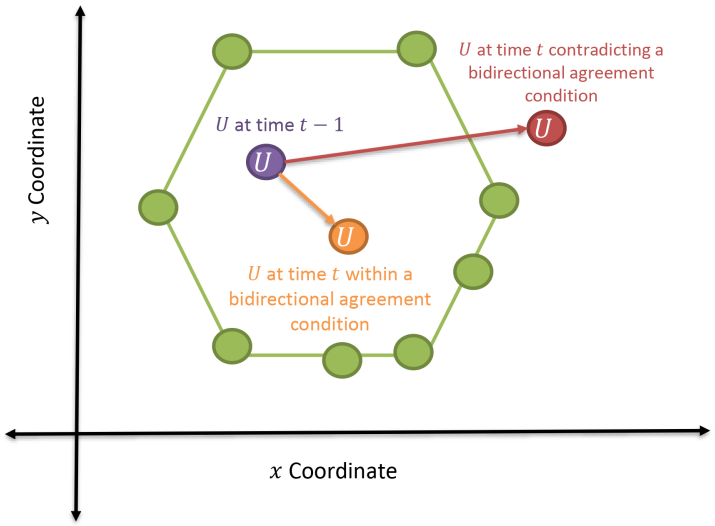


Figure 28. An example of state changing situation in the two-dimensional space. The green nodes are states of individuals at time $t-1$ and the green polygon is a neighbor-state convex hull of individual U. If U changes its state under the Bidirectional Agreement Condition, then the next state of U is always in the convex hull (orange). On the contrary, if U steps outside the convex hull (red) to make a group changes its state, then it breaks the Bidirectional Agreement Condition.

7.1.3 Leaders as state changers

When leaders initiate a group movement, if leaders are state changers, then leaders are necessary to be the first who break the Bidirectional Agreement Condition or step outside group's state convex hull. Figure 28 shows an example of state changing situation in two-dimensional state. Suppose U is a state changer and a leader, while U initiates movement, U

steps outside the group's state convex hull, which means U breaks the Bidirectional Agreement Condition. In this paper, we infer whether leaders who initiate movement are state changers by observing the association between the time that individuals break the convex hull of states (velocity, position, and direction) compared to individual leadership ranking.

7.1.4 Approach overview

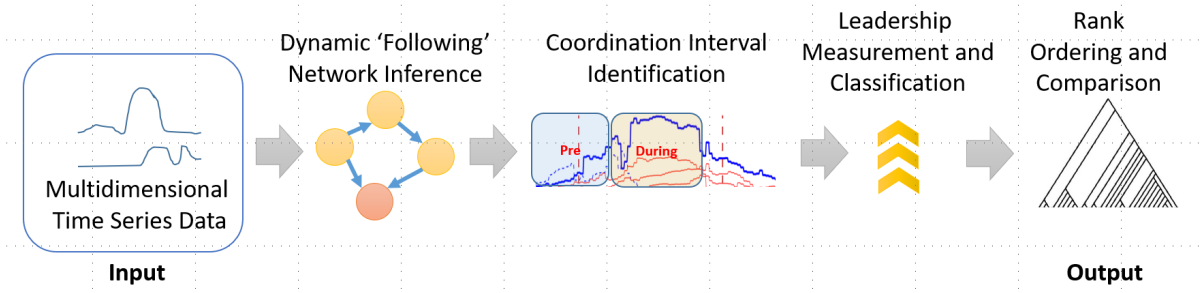


Figure 29. High-level overview of trait leadership scheme using FLICA ([1, Figure 1] used and modified with permission). An arrow between elements represents a relationship that an element at the rear of the arrow is the input of an element at the head of the arrow. For example, rank correlations are calculated by taking leadership ranking and convex hull ranking as inputs.

The high-level overview of the proposed leadership trait characterization scheme using FLICA framework [1] is shown in Figure 29. Given a time series of GPS positions of baboons,

we use FLICA to report a dynamic following network, leadership ranking, and decision-making intervals. Then, in this work, we propose measures of velocity, position, and direction convex hull containment as traits of leadership and conduct the experiments to find any significant positive/negative correlations between leadership ranking and those measures.

Let $\mathrm{D} = \{\mathrm{Q}_1, \ldots, \mathrm{Q}_n\}$ be a set of time series of positions where D consist of n time series where each $Q_i \in \mathrm{D}$ has length T (the number of time steps) and each $Q_i(t)$ is a position coordinate of the individual i at time t.

7.1.5 Leadership trait characterization scheme

7.1.5.1 The quantification of the traits of interests.

We focus on three common characterizations of a leader: being the first to move, being at the front of a group, and being the first to move in the new direction. We use the notion of the convex hull to measure the similarity of the trait value for an individual versus the group as a whole.

First, to measure the notion of being the *first to move*, we need to consider the velocity of all individuals at the previous time step. If any individual moves before others, its velocity is higher than others' velocity at the previous time step. That is, it is higher than the maximum previous velocity of any individual, or, to put it in other words, it is outside of the convex hull of the velocities in the previous time step in the positive direction (since velocity is a one-dimensional measure).

Second, to measure the notion of being at the *front of the group*, we need to consider both direction of individuals and their positions. If any individual moves toward the front of the

group, then its direction of movement is the same as the group's direction but its position is outside the group's area of the previous time step. That is, the coordinates of the individual at the front of the group are outside of the convex hull of the coordinates of the individuals in the previous time step but aligned with the direction vector of the group.

Third, to measure the notion of being the *first to move in the new direction*, we need to consider direction vectors of all the individuals. If any individual moves in the new direction, which is not the same as the group's direction, then the angle between its current direction vector and the group's direction vector at the previous time step must be high. That is, the current (angle of the) direction vector of the individual is outside the convex hull of the direction vectors of the individuals in the previous time step.

7.1.5.2 Convex hull ranking measures.

For each of the three measures we construct the convex hull in each time step and rank the individuals by the frequency with which their value in the current step is outside the convex hull of the values of all the individuals in the previous step for the same measure.

The velocity convex hull ranking score (VCH) (see Section 5.2.4.2) measures the frequency with which the discrete time series derivative (dQ_i/dt) associated with an individual i is outside the bounds of the population's (including i) discrete derivative interval in the previous time step. The highest rank of this measure indicates an individual who is the first to move in the group. Let $n \times T$ matrix VCH be a velocity convex hull score matrix where $\text{VCH}(Q_i, t) = 1$ if a time series Q_i at time t has its velocity greater than a maximum velocity of the entire group at time $t-1$ and $\text{VCH}(Q_i, t) = -1$ if Q_i has its velocity less than a minimum velocity of the

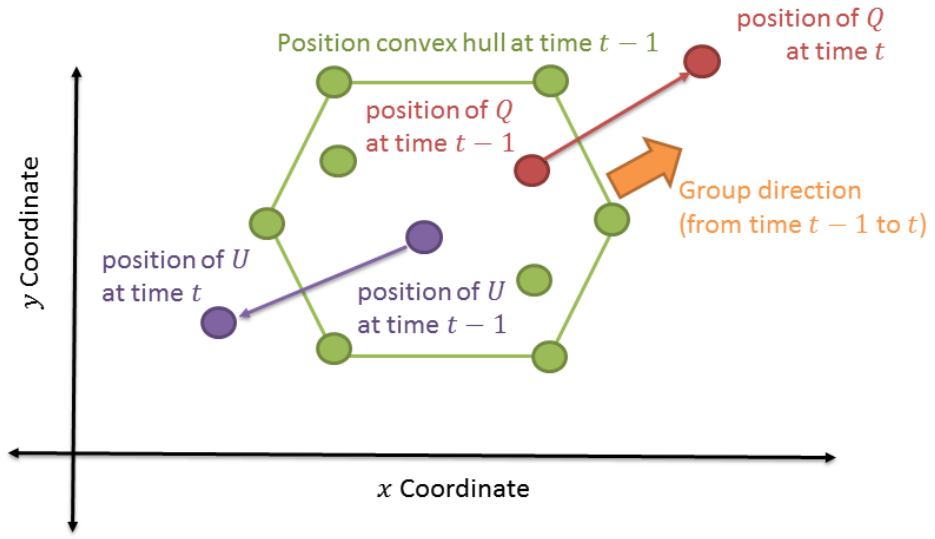


Figure 30. An example of position convex hull. Each point represents an individual and the polygon represents a convex hull boundary at time $t-1$. In this example, Q steps outside the convex hull at time t toward the group direction, while U steps outside the convex hull in the opposite direction. In this case, Q gets a score +1 and U gets a score -1 for time step t. If an individual is still in the convex hull, it gets zero score.

entire group at time $t-1$, and otherwise $\text{VCH}(Q_i, t) = 0$.

The position convex hull ranking score (PCH) (see Section 5.2.4.3) measures the frequency with which a position and direction associated with an individual i is outside the bounds of the population's position convex hull in the previous time step. A high rank of this measure indicates which individual first explores a new area before others. Let $n \times T$ matrix PCH

be a position convex hull score matrix where $\text{PCH}(Q_i, t) = 1$ if a time series Q_i at time t has its direction toward the group's direction and i's position at time t is outside the group's position convex hull at time $t-1$ (see Figure 30). In contrast, if Q_i is outside the convex hull but moving in the opposite way of group direction, then $\text{PCH}(Q_i, t) = -1$ (see Figure 30), otherwise $\text{PCH}(Q_i, t) = 0$. We consider that i is moving toward the group direction if the angle between i's direction vector and group's direction vector is between -90 and 90 degrees. Otherwise, we consider that i is moving in the opposite direction from its group movement.

The direction convex hull ranking score (DCH) measures the frequency with which the angle between individual's direction vector and group's direction vector is outside the bound of the set of angles between each individual and the group's direction vector in the previous time step. A high rank of this measure indicates an individual that frequently deviates from the group's direction of travel. Let $n \times T$ matrix DCH be a direction convex hull score matrix where $\text{DCH}(Q_i, t) = 1$ if a time series Q_i at time t has its individual-group direction angle greater than a maximum individual-group direction angle of the entire group at time $t-1$, and $\text{DCH}(Q_i, t) = -1$ if i has its individual-group direction angle lower than a minimum individual-group direction angle of the entire group at time $t-1$, and otherwise $\text{DCH}(Q_i, t) = 0$.

7.1.5.3 Rank correlation.

We deploy the Kendall rank correlation coefficient $\tau(x, y)$ [71] to infer correlation between PageRank leadership ranking (see Section 5.2.4) and the convex hull ranking.

Further, for any given threshold λ used to determine coordination events, we focus on only decision-making intervals to measure the rank correlations. We define two levels of analysis: **time-point level** and **interval level**. First, for a **time-point-level** correlation, we compute a rank correlation for each time t within any decision-making interval as follows.

Let $R_{\mathrm{PR},t} = \mathrm{argsort}(\mathrm{PR}(:,t))$ be a PageRank rank ordered list at time t such that $R(Q_i)_{\mathrm{PR},t} = q$ if an individual i is at qth rank at time t; $R(Q_i)_{\mathrm{PR},t} = 1$ if i is a leader at time t. Note that we always use argsort to represent the descending sort-order for the entire paper since the higher score implies the better rank.

$R_{\mathrm{VCH}}(t) = \mathrm{argsort}(\mathrm{VCH}(:,t))$, $R_{\mathrm{PCH}}(t) = \mathrm{argsort}(\mathrm{PCH}(:,t))$, and $R_{\mathrm{DCH}}(t) = \mathrm{argsort}(\mathrm{DCH}(:,t))$ are velocity, position, and direction convex hull rank order lists, respectively. A rank correlation between PageRank and VCH is $\tau(R_{\mathrm{PR}}(t), R_{\mathrm{VCH}}(t))$. A rank correlation between PageRank and PCH is $\tau(R_{\mathrm{PR}}(t), R_{\mathrm{PCH}}(t))$. And a rank correlation between PageRank and DCH is $\tau(R_{\mathrm{PR}}(t), R_{\mathrm{DCH}}(t))$. We define a set of time-point PageRank-VCH correlations as follows.

$$\Phi_{\mathrm{PR,VCH}} = \{\tau(R_{\mathrm{PR}}(t_1), R_{\mathrm{VCH}}(t_1)), \tau(R_{\mathrm{PR}}(t_2), R_{\mathrm{VCH}}(t_2)), \dots\} \tag{7.2}$$

Where t_i is a time point within any decision-making interval. Similarly, we can also define a set of time-point PageRank-PCH correlations and PageRank-DCH correlations in the similar way.

$$\Phi_{\mathrm{PR,PCH}} = \{\tau(R_{\mathrm{PR}}(t_1), R_{\mathrm{PCH}}(t_1)), \tau(R_{\mathrm{PR}}(t_2), R_{\mathrm{PCH}}(t_2)), \dots\} \tag{7.3}$$

$$\Phi_{\mathrm{PR,DCH}} = \{\tau(R_{\mathrm{PR}}(t_1), R_{\mathrm{DCH}}(t_1)), \tau(R_{\mathrm{PR}}(t_2), R_{\mathrm{DCH}}(t_2)), \dots\} \tag{7.4}$$

Second, for an **interval-level** rank correlation, we compute the representative correlation of the entire decision-making interval for each coordination event. Let $I = (i, j, l)$ be any coordination event, we define $\tilde{R}_{\mathrm{PR},I} = \mathrm{argsort}(\sum_{t\in[i,j]} R_{\mathrm{PR}}(t))$ as a PageRank rank ordered list during decision-making interval of coordination event I. $\tilde{R}_{\mathrm{VCH},I} = \mathrm{argsort}(\sum_{t\in[i,j]} R_{\mathrm{VCH}}(t))$, $\tilde{R}_{\mathrm{PCH},I} = \mathrm{argsort}(\sum_{t\in[i,j]} R_{\mathrm{PCH}}(t))$, and $\tilde{R}_{\mathrm{DCH},I} = \mathrm{argsort}(\sum_{t\in[i,j]} R_{\mathrm{DCH}}(t))$ are defined to be VCH, PCH, and DCH rank ordered lists of I, respectively. The PageRank-VCH rank correlation at decision-making interval of I is $\tau(\tilde{R}_{\mathrm{PR},I}, \tilde{R}_{\mathrm{VCH},I})$, the PageRank-PCH rank correlation is $\tau(\tilde{R}_{\mathrm{PR},I}, \tilde{R}_{\mathrm{PCH},I})$, and the PageRank-DCH rank correlation is $\tau(\tilde{R}_{\mathrm{PR},I}, \tilde{R}_{\mathrm{DCH},I})$. We define a set of interval PageRank-VCH correlations as follows.

$$\tilde{\Phi}_{\mathrm{PR,VCH}} = \{\tau(\tilde{R}_{\mathrm{PR},I_1}, \tilde{R}_{\mathrm{VCH},I_1}), \tau(\tilde{R}_{\mathrm{PR},I_2}, \tilde{R}_{\mathrm{VCH},I_2}), \dots\} \tag{7.5}$$

Where I_i is an ith coordination event. We can also define PageRank-PCH correlations and PageRank-DCH correlations in the similar way.

$$\tilde{\Phi}_{\mathrm{PR,PCH}} = \{\tau(\tilde{R}_{\mathrm{PR},I_1}, \tilde{R}_{\mathrm{PCH},I_1}), \tau(\tilde{R}_{\mathrm{PR},I_2}, \tilde{R}_{\mathrm{PCH},I_2}), \dots\} \tag{7.6}$$

$$\tilde{\Phi}_{\mathrm{PR,DCH}} = \{\tau(\tilde{R}_{\mathrm{PR},I_1}, \tilde{R}_{\mathrm{DCH},I_1}), \tau(\tilde{R}_{\mathrm{PR},I_2}, \tilde{R}_{\mathrm{DCH},I_2}), \dots\} \tag{7.7}$$

7.1.5.4 Leadership model classification using trait rank correlation

For model classification, we use three interval-level rank correlations from Section 7.1.5.3 as features to train a classifier. For each dataset, our framework provides a vector of features $\vec{v} = (\tau(\tilde{R}_{\mathrm{PR}}, \tilde{R}_{\mathrm{VCH}}), \tau(\tilde{R}_{\mathrm{PR}}, \tilde{R}_{\mathrm{PCH}}), \tau(\tilde{R}_{\mathrm{PR}}, \tilde{R}_{\mathrm{DCH}}))$, which represents trait characteristic of leadership model. We use Multiclass support vector machine (SVM) as our main classifier.

7.2 Experimental setup

7.2.1 Trait of leadership model

In this section, we provide three different models of traits leadership. We use these models to demonstrate that our rank correlations in Section 7.1.5.3 can be used as features to classify these models, which have different traits of leadership. All these models are in two-dimensional space. Initially, there are 20 individuals within a unit cycle. Positions of individuals are uniformly distributed within this unit cycle. Then the group moves toward a collective target.

7.2.1.1 Moving First model.

In this model, high rank individuals moves earlier than low-rank individuals. A leader moves toward target trajectory and everyone follows its hierarchy. We have ID(1) as a leader. ID(k) move first , then it is followed by ID($k+1$) with a constant time delay. The acceleration of movement for all individuals is constant. We aim to use this model as a representative model that high-rank individuals always move earlier than low-rank individuals. For this model, we set the initial velocity at 1 unit/time step and acceleration at 0.001 unit/time step^2.

7.2.1.2 Moving Front model.

This model also has an ordered hierarchy of following the same as the previous model. Nevertheless, there is no order of movement initiation. In other words, all individuals have uniformly time delay before they start moving. The group moves along a target trajectory with a constant velocity and a leader is always in the front of the group following by high-rank individuals. Lower-rank individuals follow higher-rank individuals. We aim to use this model as a representative model that high-rank individuals always explore new areas before low-rank individuals. For this model, we set the initial velocity at 1 unit/time step and acceleration at zero unit/time step^2.

7.2.1.3 Reversible Agreement model.

Compared to previous models, this model has no leader and any following hierarchy. All individuals move toward the average of group's direction with a constant velocity. This model is one of Bidirectional agreement systems that have convergence property [47]. In our case, all individual's directions converges to an average group direction, which implies the existence of coordinated movement of the group. We aim to use this model as a representative model that the group has coordinated movement without leadership hierarchy. We expect that any leadership model classification should be able to at least distinguish between leadership models and this non-leadership model. For this model, we set the initial velocity at 1 unit/time step and acceleration at zero unit/time step^2.

7.2.2 Datasets

7.2.2.1 Simulated datasets.

We create simulation datasets with the difference level of noises. We have two types of noise here: direction noise and position noise. For direction noise, instead of moving to a target direction at degree D, an individual moves toward direction $D + a$. The direction noise a is drawn randomly from normal distribution with zero mean and γ standard deviation where $\gamma \in \{0, 1, 10, 30, 60\}$. For position noise, suppose (x, y) is the next position that an individual should move to, with position noises, the actual position that the individual moves is $(x + b_1, y + b_2)$. The position noises b_1, b_2 are drawn randomly from a normal distribution with zero mean and β standard deviation where $\beta \in \{0.0001, 0.001, 0.01, 0.1, 1\}$.

For each noise setting (γ, β), we create 100 for each trait of leadership model. Each dataset contains 20 time series of individuals, which have the length as 300 time steps. In total, since we have three leadership models and 25 possible different (γ, β), we have 7,500 datasets.

7.2.2.2 Simulation datasets for degree of hierarchy structure analysis

We use simulated datasets that can be found in [1]. There are three leadership models we use in this paper: dictatorship, hierarchical model, and random model. Each model consists of 100 datasets. Each datatset has 2-dimensional time series of 20 individuals. Each time series has its length at 12000 time steps. There are 20 coordination events within each dataset.

Initially, all individuals are at their starting point. In dictatorship model (DM), a leader moves first, then everyone else follows its leader with some time delay. In hierarchical model (HM), there are four high-rank individuals, ID1, ID2, ID3, and ID4. Other none-high-rank

individuals get assigned their leaders to be one of high-rank individuals. ID1 is a global leader of all high-rank individuals that always moves first. ID2 and ID3 follow ID1 with some time delay. Then ID4 follows ID1. Lastly, the followers of ID1, ID2, ID3, and ID4 follow their leaders.

For the random model, all individuals move together toward a target direction. However, these individuals never follow any specific individuals. Hence, there are coordination events in this model but there are no leaders.

7.2.2.3 Baboon trajectories.

The data consists of time series of latitude-longitude location pairs for each baboon every second. 16 individuals whose collars remained functional throughout the time are analyzed for a case study of a merging coordination event. See Section 4.2.1 for more information.

7.2.2.4 Fish schools trajectories.

The fish dataset is a set of time series of fish positions from a video record of a school of golden shiners (*Notemigonus crysoleucas*). See Section 4.2.2 for more information.

7.2.3 Sensitivity analysis in model classification

We separate simulation dataset into the groups based on the value of noise setting (γ, β). For each group, it consists of 100 datasets of Moving First model, 100 datasets of Moving Front model, and 100 datasets of Reversible Agreement model. We report 10-fold Cross validation of model classification for each group of datasets having the same noise level (γ, β). We also report the rank correlation between the ground-truth leadership rank vs. inferred leadership

TABLE XIV. Description of hypothesis tests used in this paper. A significant level has been set at $\alpha = 0.001$ for all experiments.

	Method	**Null hypothesis H_0**
Zero mean/median test	t-test	A sample has a normal distribution with zero mean and unknown variance.
	Sign test	A sample has a distribution with zero median.
	Wilcoxon signed rank test	A sample has a symmetric distribution around zero median.
Normality test	Kolmogorov-Smirnov test	A sample comes from a normal distribution.
	Chi-square goodness-of-fit test	A sample comes from a normal distribution with a mean and variance estimated from a sample itself.
	Jarque-Bera test	A sample comes from a normal distribution with an unknown mean and variance.
	Anderson-Darling test	A sample comes from a normal distribution.

rank from our framework to measure the ability of leadership inference within difference level of noises.

7.2.4 Hypotheses tests

In this section, we aim to design a hypothesis testing scheme to address three hypotheses: (1) individuals who act as leaders (identified by FLICA) are individuals who move first, initiating their movements before others in their group in the decision-making period prior to coordinated movement, (2) individuals who act as leaders are individuals who always explore

a new area before others prior to a coordinated movement, and (3) individuals who act as leaders are individuals who always align with the group direction. We define leaders as individuals who possess a highly ranked position in a PageRank rank ordered list. The hypothesis testing methods we used can be categorized into two categories: zero mean/median test and normality test. For the zero mean/median test, we aim to test whether a positive/negative correlation exists between PageRank and convex hull ranking in both time-point and interval levels. For normality tests, we aimed to determine whether correlation samples come from a normal distribution. If not, the interpretation of tests which assume normally distributed data, *e.g., t*-test, should be considered carefully. The full list of hypothesis testing methods we used is in Table XXVI. We set significant level at $\alpha = 0.001$ for all tests.

7.2.5 Parameter setting

For simulation dataset, we set the time window $\omega = 60$ and $\delta = 6$, which is the optimal setting since the simulation dataset has time delay less than 5 time steps by design. For the analysis in baboon dataset, we set the time window $\omega = 240$ and $\delta = 24$. For the fish dataset, the time window $\omega = 285$ and $\delta = 28$. Both parameter settings of baboon and fish datasets are set based on the fact that these setting can infer the highest number of following relations per following group on average. The time sliding window parameter δ serves to trade off computation versus the sampling rate of the time series process. The FLICA has time complexity $\mathcal{O}(n^2 \times t \times \omega)$. The network density decision-making threshold λ was set at 25th 50th, 75th, and 99th percentile of network density values for the baboon dataset to detect

decision-making intervals. For the simulation and fish datasets, we already have the decision-making intervals so we do not need to set λ.

7.3 Results

7.3.1 Traits of leader classification: sensitivity analysis

Figure 31 shows the result of sensitivity analysis in the model classification. Loss values of 10-fold cross validation are in the top of figure. A loss value is a percentage of datasets that the classifier predicted them into wrong classes. Figure 31 (top) shows that when the level of noises increase, classifier produces more errors. According to the cross validation result, our framework can distinguish between leadership models (Moving first and Moving front models) and non-leadership model (Reversible Agreement Model). Additionally, the result suggests that position noise affected the classification result than direction noise. When the position noise level reach at 1, which is the diameter of group movement, the leadership rank is less consistent with the ground-truth rank (Figure 31 bottom). This indicates that both leadership ranking and traits of leadership inference are hard to perform under high-level of position noise. In general, this result shows that our framework performs accurately even if an input data is noisy until a certain degree of noises.

7.3.2 Traits identification of baboons movement

The distributions of rank correlations inferred from the baboon dataset are in Figure 32. At the time-point level, the distribution of PageRank and velocity convex hull (PR-VCH) correlation $\Phi_{\mathrm{PR,VCH}}$ is at the top-left of the figure, the distribution of PageRank and position convex hull (PR-PCH) correlation $\Phi_{\mathrm{PR,PCH}}$ is at the top-middle and the distribution of PageRank and

TABLE XV

PR-VCH, PR-PCH, and PR-DCH rank correlations from the baboon dataset under different thresholds.

		Time-point level		Interval level	
	Percentile	**Mean**	STD.	**Mean**	STD.
PR-VCH Corr.	25th	0.03	0.20	-0.09	0.18
	50th	0.03	0.19	-0.07	0.18
	75th	0.03	0.19	-0.06	0.19
	99th	0.03	0.19	-0.07	0.21
PR-PCH Corr.	25th	0.00	0.23	0.09	0.20
	50th	0.01	0.25	**0.12**	0.21
	75th	0.06	0.27	**0.18**	0.22
	99th	**0.15**	0.30	**0.32**	0.22
PR-DCH Corr.	25th	-0.0082	0.1739	0.03	0.22
	50th	-0.0098	0.1745	-0.01	0.23
	75th	-0.0251	0.1818	**-0.09**	0.25
	99th	-0.0552	0.1827	**-0.24**	0.24

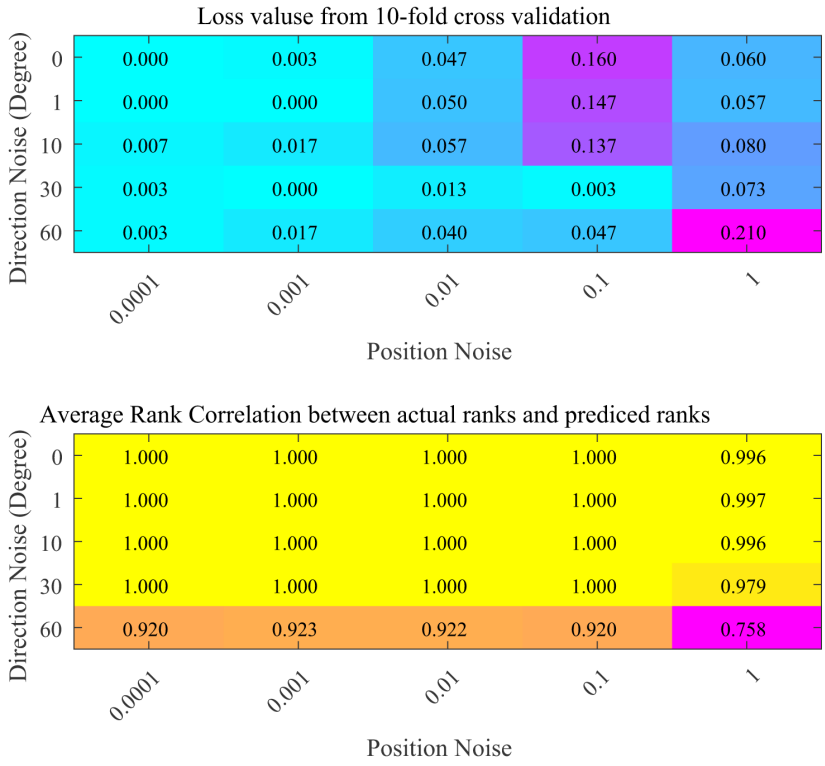


Figure 31. Sensitivity analysis in model classification task from simulation datasets with different noise levels. (Top) 10-fold cross validation loss values. Each element in the table represents the loss value of each each noise setting (γ, β). (Buttom) Rank Correlation between actual leadership ranking vs. predicted ranking from moving first and moving front models.

direction convex hull (PR-DCH) $\Phi_{\text{PR,DCH}}$ is at the top-right of the figure. For the interval-level correlations, the distribution $\tilde{\Phi}_{\text{PR,VCH}}$ is at the bottom-left, the distribution $\tilde{\Phi}_{\text{PR,PCH}}$ is at the the bottom-middle, and the distribution $\tilde{\Phi}_{\text{PR,DCH}}$ is at the the bottom-right. Table XV illustrates the means and standard deviations of these correlation distributions.

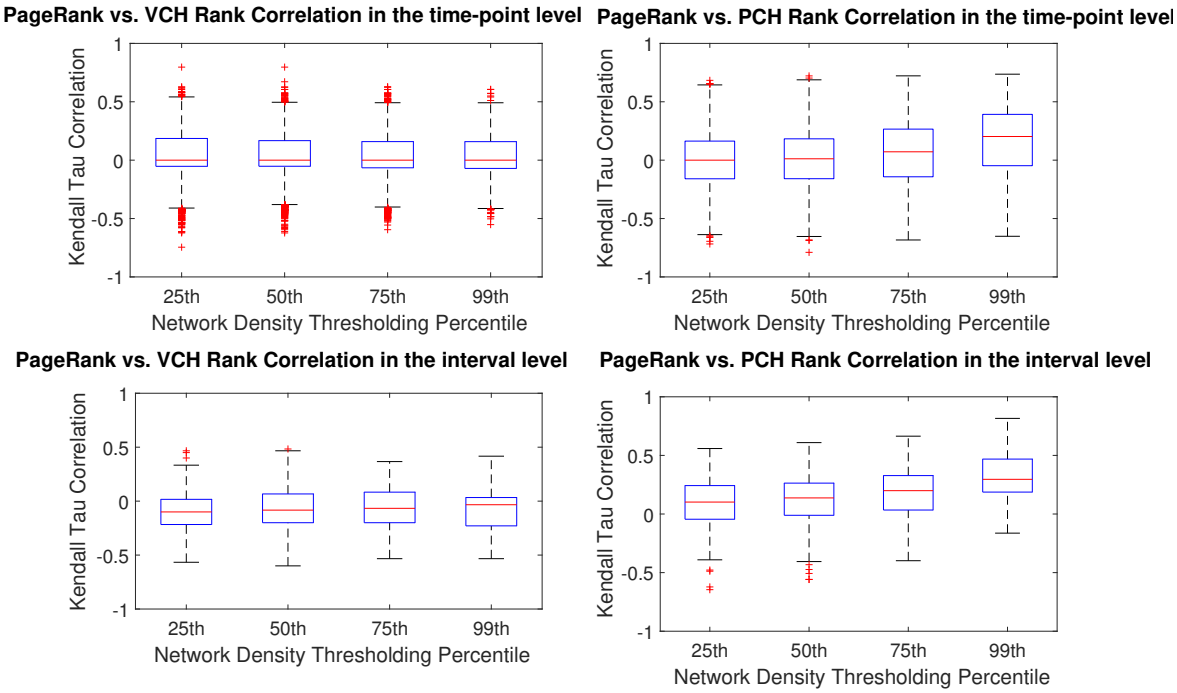


Figure 32. Comparison of PR-VCH, PR-PCH, and PR-DCH rank correlations under different thresholds. For PR-VCH correlation, the results in both time-points level (Top-left) and interval level (Bottom-left) show that there are no strong correlations between leadership and VCH ranking. In contrast, leadership and PCH rankings have positive correlations in both time-points level (Top-middle) and interval level (bottom-middle), as well as PR-DCH correlation has a negative correlation at the interval level (bottom-right).

Figure 32 and Table XV suggest that there is no correlation between PageRank and velocity convex hull ranking at both the time-point level (Figure 32's top-left) and the interval

TABLE XVI

Hypothesis test results of PR-PCH and PR-DCH correlation. The zero value implies that a test fails to reject H_0 while one implies a test sucessfully rejects H_0 with $\alpha = 0.001$.

		PR-PCH Corr. at a time-point level				**PR-PCH Corr. at an interval level**				**PR-DCH Corr. at an interval level**			
	Tests/λ THS	25th	50th	75th	99th	25th	50th	75th	99th	25th	50th	75th	99th
Zero mean/median test	ttest	0	1	1	1	1	1	1	1	0	0	1	1
	Sign test	0	1	1	1	1	1	1	1	0	0	1	1
	Wilcoxon test	0	1	1	1	1	1	1	1	0	0	1	1
Normality test	KS test	1	1	1	1	1	1	1	1	1	1	1	1
	Chi-square test	1	1	1	1	0	0	0	0	0	0	0	0
	Jarque-Bera test	1	1	1	1	0	0	0	0	0	0	0	0
	Anderson-Darling test	1	1	1	1	0	0	0	0	0	0	0	0

level (Figure 32's bottom-left). In contrast, positive correlations exist between PageRank and position convex hull in both levels (Figure 32, top and bottom middle). Moreover, negative correlations exist between PageRank and direction convex hull in interval level (Figure 32, bottom right). When we set a higher percentile threshold, we get stronger coordination events; a stronger coordination event has a higher number of following relations. Both Figure 32 and Table XV illustrate that the rank correlations between PageRank and position convex hull ranking are higher, while the rank correlations between PageRank and direction convex hull ranking are lower when we set a stronger threshold. However, there is not a large difference in the correlations of PageRank and velocity convex hull when we varied the threshold value.

Based on this result, due to weak correlations of PR-VCH in both time-point and interval levels and PR-DCH in the time-point level, we decided to conduct the hypothesis tests only in the PR-PCH rank correlation samples in both levels while conducting the hypothesis tests for PR-DCH in the interval level.

The result of these hypotheses tests are shown at Table XVI. In the aspect of normality test results, correlations at time-point level of PCH are less normal compared to the PCH's and DCH's correlations at the interval level. This implies that the result of t-test at the PCH's time-point level should be interpreted carefully.

In the aspect of zero/mean median hypothesis test, with the significant level at $\alpha = 0.001$, PageRank and position convex hull ranking have positive correlations far from zero. This implies that individuals who act as leaders tend to explore new areas before other individuals during decision-making intervals.

In contrast, PageRank and direction convex hull ranking has negative correlations far from zero at the 75th and 99th percentile thresholds. This implies that individuals who act as leaders tend to align with the group's direction (or, more intuitively, the group is aligned with the leader's direction) while non-leaders frequently attempt to change the direction but nobody follows. In other words, high-rank individuals control the group direction and this is why they are almost always inside the direction convex hull. When high-rank individuals move in any given direction, the group follows almost immediately and this makes the group's direction the same as the leading individuals' direction.

TABLE XVII

Normal confidence intervals of PR-PCH and PR-DCH correlations from the baboon dataset at the interval level with $\alpha = 0.001$

		Normal Confidence Interval			
		Mean μ		STD.	
	Percentile Threshold	Lower bound	Upper bound	Lower bound	Upper bound
PR-PCH	25th	0.06	0.13	0.18	0.23
	50th	0.08	0.16	0.18	0.23
	75th	0.13	0.23	0.19	0.25
	99th	0.19	0.44	0.16	0.35
PR-DCH	25th	-0.01	0.07	0.20	0.25
	50th	-0.05	0.03	0.21	0.27
	75th	-0.14	-0.04	0.21	0.29
	99th	-0.38	-0.11	0.17	0.37

Finally, for interval-level correlations of PR-PCH and PR-DCH ranking, we reported the normal confidence intervals at Table XVII. We only reported the confidence intervals of interval-level correlations because of the normality test results; time-point level of PR-PCH correlation distributions seem not to be normal (see Table XVI) while the rest of the cases are normal. All normal confidence intervals of PR-PCH correlation distributions have their lower bound greater than zero, while the upper bounds of PR-DCH correlation at the 75th and 99th percentile thresholds are below zero. This supports the hypotheses that there exist a positive correlation between PageRank and PCH ranking and a negative correlation between PageRank and DCH at the interval level.

7.3.3 Traits identification of fish movement

The distributions of rank correlations inferred from the fish dataset are in Figure 33. At the time-point level, the distribution of $\Phi_{\text{PR,VCH}}$, $\Phi_{\text{PR,PCH}}$, and $\Phi_{\text{PR,DCH}}$ are at the left of the figure, while the right of the figure contains rank correlations at the interval level. Table XVIII illustrates the means and standard deviations of these correlation distributions.

At the time-point level, Figure 33 and Table XVIII suggest that there is no correlation between PageRank vs. VCH ranking and PageRank vs. DCH ranking, while we have positive rank correlations of PageRank vs. PCH on average. At the interval level, $\tilde{\Phi}_{\text{PR,VCH}}$ and $\tilde{\Phi}_{\text{PR,PCH}}$ have positive values on average, while $\tilde{\Phi}_{\text{PR,DCH}}$ has negative values on average.

Based on this result, due to the weak correlations of PR-VCH and PR-DCH in a time-point level, we decided to conduct the hypothesis tests only in the PR-PCH rank correlation samples

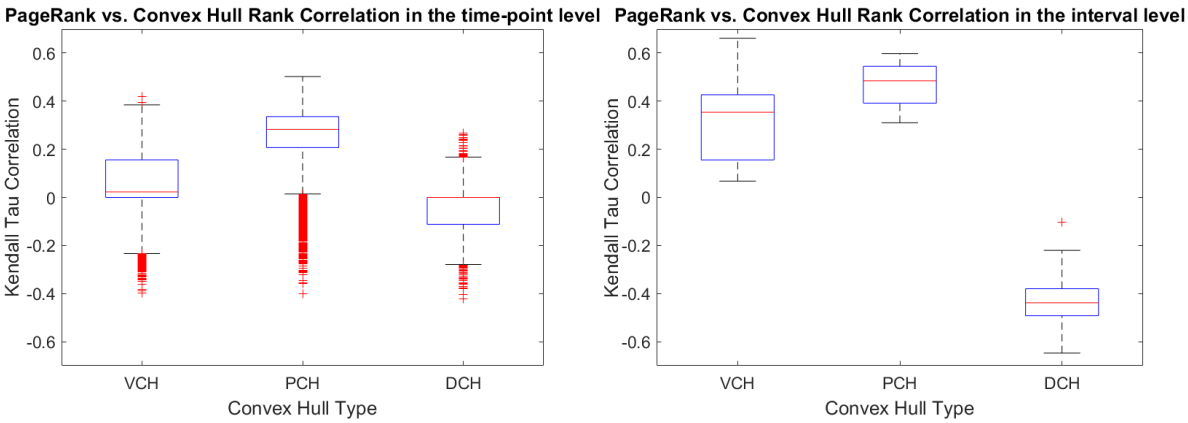


Figure 33. Comparison of PR-VCH, PR-PCH, and PR-DCH rank correlations in both time-point and interval levels from the fish movement dataset. In the time-point level (left), the result shows that leadership vs. VCH, and leadership vs. PCH rankings have positive correlations, while leadership and DCH has negative correlation. In the interval level (right), leadership vs. VCH and leadership vs. PCH rankings have stronger positive correlations than time-point level, while leadership and DCH also has stronger negative correlation.

in both levels while conducting the hypothesis tests for PR-VCH and PR-DCH in the interval level.

The result of hypotheses tests from the fish dataset are shown at Table XIX. In the aspect of normality test results, correlations at time-point level of PCH are less normal compared to the VCH's, PCH's and DCH's correlations at the interval level. This implies that the result of t-test at the PCH's time-point level should be interpreted carefully.

TABLE XVIII

PR-VCH, PR-PCH, and PR-DCH rank correlations from the fish dataset.

	Time-point level		Interval level	
	Mean	Std.	**Mean**	Std.
PR-VCH Corr.	0.05	0.12	0.32	0.16
PR-PCH Corr.	0.26	0.12	0.47	0.09
PR-DCH Corr.	-0.05	0.08	-0.43	0.12

In the aspect of zero/mean median hypothesis test, with the significant level at $\alpha = 0.001$, both PageRank vs. velocity convex hull ranking and PageRank vs. position convex hull ranking have positive correlations far from zero on average. The result demonstrates that individuals who act as trained fish tend to move earlier and explore new areas before other individuals during coordination events. On the contrary, PageRank vs. Direction convex hull ranking has negative correlations far from zero on average. The result implies that when trained fish moves in any given direction, the group follows almost immediately and this makes the group's direction the same as a trained fish's direction.

We also reported the normal confidence intervals at Table XX. We only reported the confidence intervals of interval-level correlations because of the normality test results; time-point

TABLE XIX

Hypothesis test results of PR-VCH, PR-PCH, and PR-DCH correlations in time-point level and interval level from the fish movement dataset. The zero value implies that a test fails to reject H_0 while one implies a test successfully rejects H_0 with $\alpha = 0.001$.

		Time-point level	Interval level		
		PR-PCH Corr.	PR-VCH Corr.	PR-PCH Corr.	PR-DCH Corr.
Zero mean/median test	ttest	1	1	1	1
	Sign test	1	1	1	1
	Wilcoxon signed rank test	1	1	1	1
Normality test	Kolmogorov-Smirnov test	1	1	1	1
	Chi-square goodness-of-fit test	1	0	0	0
	Jarque-Bera test	1	0	0	0
	Anderson-Darling test	1	0	0	0

level of PR-PCH correlation distributions seem not to be normal (see Table XIX) while the rest of the cases are normal.

According to Table XX, the normal confidence intervals of PR-VCH and PR-PCH correlation distributions have their lower bound greater than zero, while the upper bound of PR-DCH correlation at is below zero. This supports the hypotheses that there exist a positive correlation between PageRank and VCH ranking as well as PageRank and PCH ranking, while there is a negative correlation between PageRank and DCH at the interval level.

TABLE XX

Normal confidence intervals of PR-VCH, PR-PCH. and PR-DCH correlations from the fish dataset at the interval level with $\alpha = 0.001$

	Normal Confidence Interval			
	Mean μ		STD.	
	Lower bound	Upper bound	Lower bound	Upper bound
PR-VCH	0.20	0.45	0.11	0.30
PR-PCH	0.41	0.54	0.06	0.16
PR-DCH	-0.52	-0.34	0.08	0.21

7.3.4 Traits of leaders as measure of degree of hierarchy structure

Another application of trait-rank correlations we proposed here is to use these correlations to measure the degree of hierarchy structure in the datasets. The hierarchy structure here is the order of early movement. If the datasets contain a high degree of order of movement, then some specific individuals (e.g. high-rank individuals) always move before other individuals. In contrast, if datasets contains no order of movement, then there is no specific order of individuals who move before others. Figure 34 illustrates the distributions of PR-VCH rank correlations of datasets (Section 7.2.2.2) from three leadership models. As we expected, since hierarchical

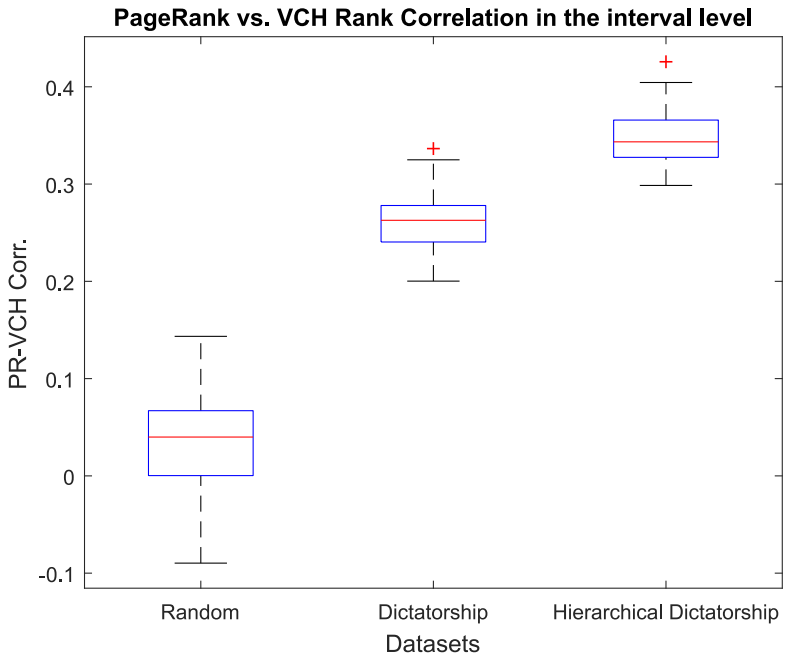


Figure 34. The distributions of PR-VCH rank correlations of datasets from three leadership models. The higher rank correlation implies the higher degree of hierarchy structure of early movement order in the model.

model has a higher degree of structure than the dictatorship model, hence, it has the highest value of PR-VCH rank correlations. The dictatorship model has the second highest value of PR-VCH rank correlations since there is a weak order of early movement; a leader always moves first. Lastly, the random model has the PR-VCH rank correlations around zero since it has no order of early movement.

In conclusion, the result implies that the higher rank correlation implies the higher degree of hierarchy structure of early movement order in the model.

Table XXI shows the mean and standard deviation of PR-VCH rank correlations from both simulated and biological datasets. The result shows that baboon datasets have PR-VCH rank correlations nearly zero in all threshold of coordination events, while fish datasets have PR-VCH rank correlations nearly the Hierarchical model. This implies that baboons have no hierarchy of early movement while schools of fish have pretty high degree of movement order.

7.4 Conclusions

In this chapter, we proposed a framework for testing the correspondence between behavioral traits and leader individuals in the context of movement initiation. We focused on three hypotheses. First, individuals who act as leaders tend to move before others in their group in the period preceding coordinated movement. Second, individuals who act as leaders tend to move into new areas before others prior to a coordinated movement. Third, individuals who act as leaders tend to set the group's direction of travel. We constructed a dynamic following network and used the simple notion of convex hull as the measure of degree of difference of the velocity, position, and direction of an individual from its group. We use proposed traits of leaders for model classification. We evaluated the classification task on simulated movement data. We tested our proposed approach in baboon movement and fish movement datasets using the time-series leadership inference framework, FLICA.

We found that during baboon decision-making intervals before a period of coordinated troop movement, there was a positive correlation between an individual's leadership ranking and the

TABLE XXI

The means and standard deviations of PR-VCH rank correlations from both simulated and biological datasets.

Datasets	**Mean**	**STD.**
Hierarchical model	0.35	0.03
Dictatorship model	0.26	0.03
Random	0.04	0.05
Baboon (25th-THS)	-0.09	0.18
Baboon, (50th-THS)	-0.07	0.18
Baboon (75th-THS)	-0.06	0.19
Baboon (99th-THS)	-0.07	0.21
Fish	0.32	0.16

frequency with which an individual decided to step outside the group to explore a new area. Moreover, there was a negative correlation between leadership ranking and the frequency which individuals misaligned with the group's direction. We drew this conclusion from the hypothesis testing of the distribution of correlations between leadership ranking and convex hull measures, constructed by the proposed framework. However, there was no strong correlation between the frequency of early movement and leadership ranking.

In the fish dataset, we found that there were a positive correlation between the leadership ranking and the order of movement ranking, as well as leadership vs. order of exploring new areas ranking. On the contrary, on average, there was a negative correlation between the leadership ranking and the frequency which individuals misaligned with the group's direction. These results suggest that trained fish seems to move earlier than others fish to the new area and the untrained fish align with trained fish quickly.

Our work establishes a general framework to draw conclusions about leadership characteristics of individuals initiating movement and to test long standing common assumptions about the behavioral traits the leaders possess. Our framework is sufficiently general to be applied to any movement dataset and any set of traits directly computable from the data. In the next chapter, we will introduce the proposed framework that can be used to infer a coordination strategy of each individual that can make the entire group achieve coordination.

CHAPTER 8

INFERRING COORDINATION MECHANISMS FROM TIME SERIES OF MOVEMENT DATA

8.1 Introduction

COORDINATION STRATEGY INFERENCE PROBLEM: To reach a group consensus, individuals have to coordinate with others. There are many strategies each individual can use to achieve coordination at the group level. **Given time series of individual activities and a set of candidate strategies, the goal is to find the set of original strategies individuals used that lead to the group consensus.**

How do groups of individuals achieve consensus in movement decisions? Do people follow their friends or the one predetermined leader? Do baboons have a hierarchy which determines the order of movement? Is the path of a fish determined by its nearest neighbors? To address these questions computationally, in Section 3.5, we formalized COORDINATION STRATEGY INFERENCE PROBLEM. In this setting, a group of multiple individuals moves in a coordinated manner towards a target path. Each individual uses a specific strategy to follow others (e.g. nearest neighbors, pre-defined leaders, preferred friends). In this chapter, we propose two coordination strategies: hierarchy (e.g. pre-defined leaders) and non-hierarchy (e.g. nearest neighbors) models (Section 8.2). In Section 8.3, we analyze a convergence property of these

strategies and other properties. Given a time series that include coordinated movement and a set of candidate strategies as inputs, we provide (the first, to the best of our knowledge) methodology to infer the set of strategies that each individual uses to achieve movement coordination at the group level. We evaluate and demonstrate the performance of the proposed framework by predicting the direction of movement of an individual in a group in both simulated datasets as well as two real-world datasets: a school of fish and a troop of baboons (Section 8.4). The results in Section 8.5 show that our approach is highly accurate in inferring the correct strategy in simulated datasets even in complicated mixed strategy setting. Animal data experiments show that fish, unsurprisingly, follow their neighbors, while baboons have a preference to follow specific individuals. Our methodology easily generalizes to arbitrary time series data, beyond movement data.

In the next section, we introduce a concept of convergence in multi-agent systems and the relationship between convergence and coordination strategy.

8.2 Models and properties

As a reminder, we use the following notation throughout Chapter **??**:

- $\mathcal{N} = \{1, \ldots, n\}$ is a set of agents.
- $\mathcal{I} \subseteq \mathcal{N}$ is a set of informed agents.
- $S_i(t)$ is a state value of agent i at time t, where $S_i(t) \in \mathbb{R}^d$.
- $S^t = \{S_i(t)\}$ is a set of individual states at at time t.
- $S_i = (S_i(0), \ldots, S_i(T))$ is state time series of agent i.
- $S_w = (S_w(0), \ldots, S_w(T))$ is a target path where $S_w(t) \in \mathbb{R}^d$ is a target state at time t.

- $\mathcal{H} = \{h_i\}$ is a set of strategy functions that agents use to update their current state where $h_i : \mathbb{R}^d \to \mathbb{R}^d$.
- $\mathcal{S} = \{S_i\}$ is a set of state time series generated by agents using some set of strategy functions $\mathcal{F} \subseteq \mathcal{H}$.
- $\sigma \in [0, 1]$ is a noise-tolerance threshold.

For the convergence of multi-agent systems, we adopt a notion of ϵ-convergence from [47].

Definition 19 (ϵ-convergence) *Given $S^0 = \{S_1(0), \ldots, S_n(0)\}$, a system, which is a set of strategy functions, is said to ϵ-converge if, for $0 < \epsilon \leq 1/2$, there exists a time constant $t_c > 0$ such that for all $t > t_c$, a set of n agent's states $S^t = \{S_1(t), \ldots, S_n(t)\}$ can be partitioned into disjoint subsets, where the maximum distance between any pair of agents' states $S_i(t), S_j(t)$ from the same subset is less than or equal ϵ.*

Definition 20 (ϵ-convergence of time series) *Given two time series S_1, S_2, we say that S_1 ϵ-converges toward S_2 at time t if, for all time $t' \geq t$, the distance between $S_1(t')$ and $S_2(t')$ is less than or equal ϵ, where $0 < \epsilon \leq 1/2$.*

Now, we can state the relationship between ϵ-convergence and coordination strategy.

Proposition 8.2.1 *If all time series generated by a set of strategy functions $\mathcal{F} \subseteq \mathcal{H}$ ϵ-converge toward a target path S_w, then $\mathcal{F}$ is a set of coordination strategies, where $\sigma = 1 - \epsilon$.*

Proof Suppose all time series generated by a set of strategy functions $\mathcal{F} \subseteq \mathcal{H}$ ϵ-converge toward a target path S_w. At the converging time $t \in [t_1, \ldots]$ every agent's state is within its group convex hull centered at $S_w(t)$ that has the diameter at most ϵ. For some time $t_2 \geq t_1$, every

time series has a distance between each other at most ϵ. By setting $\sigma = 1 - \epsilon$, this implies that every time series σ-follows time series S_w. By assigning all agents that have the state time series the same as S_w to be informed agents, since others follow S_w with some time delay, therefore, we have the $1-\epsilon$-coordination interval $[t_2, \dots]$ and all informed agents are initiators.

Proposition 8.2.2 *Let $\mathcal{H} = \{h_k\}$ be a set of pure strategy functions. If all agents use any $h_i \in \mathcal{H}$ as a pure strategy function and their state time series ϵ-converge toward a target path S_w, then a mixed strategy function f', created by a linear combination of functions in $\mathcal{H}$, generates a time series that ϵ-converges toward S_w.*

Proof Suppose all functions in $\mathcal{H}$ generate state time series that ϵ-converge toward S_w. At the equilibrium time t, when all strategies converge, any strategy in $\mathcal{H}$ that agent i uses ensures that i's state $S_i(t)$ is in the convex hull of states centered at $S_w(t)$ and has a diameter at most ϵ, since a linear combination of values within a convex hull is still in a convex hull. Therefore, a mixed strategy function f' that is created by a linear combination of functions in $\mathcal{H}$ generates a time series that ϵ-converges toward S_w.

8.2.1 Convergence models

8.2.1.1 Hierarchical Model Dynamic System (HM)

Let L be an informed agent. Let a directed acyclic graph (DAG) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be graph, where $\mathcal{V}$ is a set of agent nodes and $\mathcal{E}$ is a set of probabilistic edges, so that if $p_{i,j}$ is a probability that i follows j, then $(i, j) \in \mathcal{E}$ has the weight $p_{i,j}$. We call $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ a probabilistic following network.

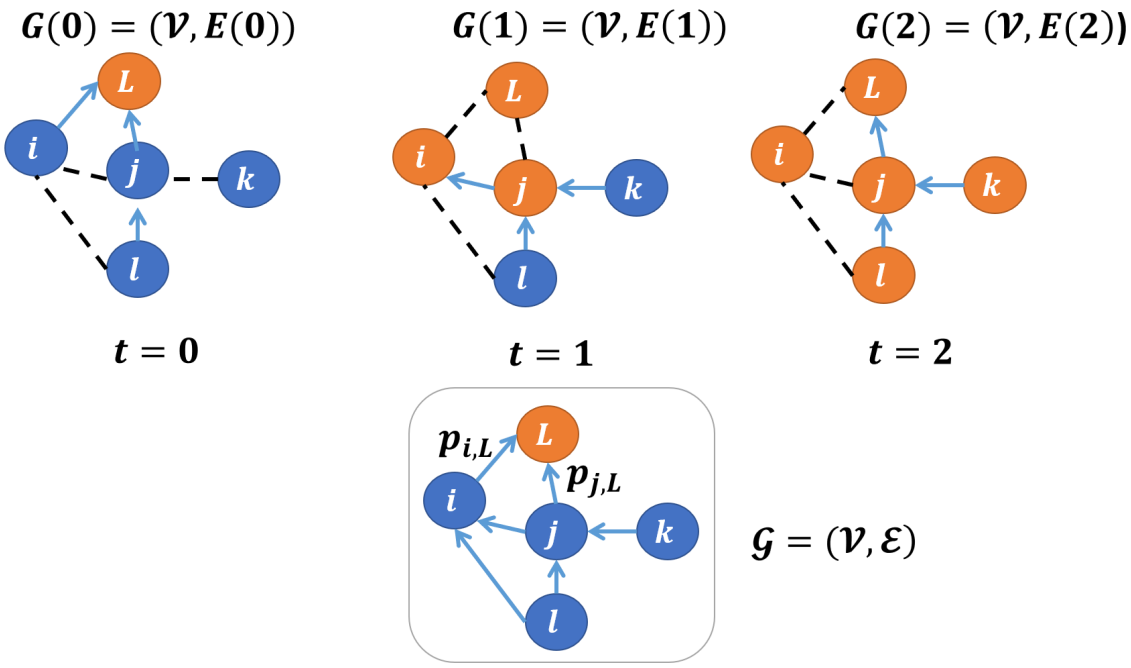


Figure 35. An example of communication networks $G(t) = (\mathcal{V}, E(t))$ between $t = 0$ and $t = 3$ (above). These networks are the realization of the probabilistic following network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ (below). The arrows represent the directed edges while the dashed lines are empty edges. When the time step increases, the informed agent L can increasingly spread its state (orange node) to more follower nodes (blue nodes).

In this model, $\mathcal{G}$ is connected and every node has a path to a leader node L. For every time step $t \geq 0$, the system generates a communication network $G(t) = (\mathcal{V}, E(t))$, which is a realization of $\mathcal{G}$. The example of the process of generation of a communication network is shown in Figure 35.

Let $S^0 = \{S_1(0), \ldots, S_n(0)\}$ be a set of agent's initial states, $\mathcal{N}_i^t = \{j|(i,j) \in E(t)\} \cup \{i\}$ be a set of neighbors of i in $G(t)$ that i follows, and S_w be a target path. At any time t, the informed agent L updates its state to be $S_w(t)$. For any other uninformed agent i, it updates the state $S_i(t)$ according to the aggregation of its neighbors' states. Formally, we have a strategy function for this model as follows:

$$f_{\mathrm{HM}}(S^{t-1}, i) = \begin{cases} S_w(t), & \text{if } i = L \\ \frac{1}{|\mathcal{N}_i^t|} \sum_{j \in \mathcal{N}_i^t} S_j(t-1), & \text{otherwise.} \end{cases} \tag{8.1}$$

Agents use the above strategy function to update the state $S_i(t) = f_{\mathrm{HM}}(S^{t-1}, i)$ in this model. In cooperative control literature, the Equation 8.1 is called a local voting protocol [45]. A system is known to converge if each communication network $G(t)$ stays the same all the time and has a spanning tree that has a leader node L as the root [45]. This is why $\mathcal{G}$ must be connected in order to make a system converge.

Theorem 8.2.3 *Let $S^0 = \{S_1(0), \ldots, S_n(0)\}$ be a set of agents' initial states. If all agents use HM strategy (Equation 8.1) to update their states, then all agents' state time series ϵ-converge toward a target state $S_w(t)$ with the expectation of the convergence time at most $t_c = n \cdot \max_i(log_2(\frac{DIST(S_i(0), S_w(0))}{\epsilon})/p^*)$ time steps if $S_w(t) = S_w(0)$ for all $t > 0$ and $p* = \min_{k,l \in \mathcal{N}, p_{k,l} > 0} p_{k,l}$.*

Proof In the first time step, $S^0 = \{S_1(0), \ldots, S_n(0)\}$ forms a convex hull and $S_w(0)$ is inside this convex hull because $S_w(0) \in S^0$. For any agent $i \in \mathcal{N}_L^t$, according to Equation 8.1, the distance $\mathrm{DIST}(S_L(t), S_i(t))$ reduces by half whenever the link $(i, L) \in E(t)$. Let $T_i \sim$

Binomial$(t_i, p_{L,i})$ be a random variable of the number of steps it takes until the appearance of a link (i, L) such that $\mathrm{DIST}(S_L(t), S_i(t)) = \epsilon$, using t_i trials. We can find the expectation of the time $\mathbb{E}(T_i)$, *i.e.,* the expected number of trials $\hat{t}_i$ until $\mathrm{DIST}(S_L(t), S_i(t)) = \epsilon$.

From Equation 8.1,

$$\epsilon = \frac{\mathrm{DIST}(S_i(0), S_w(0))}{2^{T_i}}.$$

Then, by definition of the Binomial expectation,

$$\mathbb{E}(T_i) = \hat{t}_i \times p_{i,L} = \log_2 \left(\frac{\mathrm{DIST}(S_i(0), S_w(0))}{\epsilon} \right).$$

Therefore,

$$\hat{t}_i = \frac{1}{p_{i,L}} \log_2 \left(\frac{\mathrm{DIST}(S_i(0), S_w(0))}{\epsilon} \right).$$

In general, we can have an upper bound $t_c \geq \hat{t}_i$ of the expectation of the convergence time as follows:

$$t_c = n \cdot \max_i \left\{ \frac{1}{p^*} \log_2 \left(\frac{\mathrm{DIST}(S_i(0), S_w(0))}{\epsilon} \right) \right\},$$

where

$$p* = \min_{k,l \in \mathcal{N}, p_{k,l} > 0} p_{k,l}.$$

According to Theorem 8.2.3 and Proposition 8.2.1, if the target path S_w has its target state $S_w(t)$ as a fixed point: $S_w(t) = S_w(0)$ for all $t > 0$, then the set of strategy functions $\mathcal{F}$ that contains only HM strategy functions is a set of coordination strategies. In other words, if all agents use f_{HM} to update their states, then their states converge to a target path. Therefore,

a coordination interval exists in their state time series. In contrast, if a target state $S_w(t)$ can be changed, the the group still follows the path S_w, because only L influences the group and L's state path is S_w. However, the convergence might not exist if the difference between two consecutive time steps within the target path is always greater than the group convergent rate.

8.2.1.2 Local Reversible Agreement system (LRA)

Let $P^0 = \{P_1(0), \ldots, P_n(0)\}$ be a set of physical points, S^0 be a set of initial states, S_w be a target path, L be an informed agent who updates its state in correspondence to S_w, and $g(P^t, S^t, i)$ be a projection function that agents use to update their physical points. If a state point is a velocity vector, then the projection function is simply the current position plus the velocity vector. First, for $t > 0$, we update the physical point $P_i(t) = g(P^{t-1}, S^{t-1}, i)$. Second, we create a set of Delaunay triangulations from P^t to create a communication network $G(t) = (\mathcal{V}, E(t))$. If $P_i(t)$ and $P_j(t)$ form the same triangle within the physical space, then $(i, j) \in E(t)$. Third, we update a state of each agent based on the structure of $G(t)$. The example of how to find the neighbors of each individual in LRA is in Figure 36, which defines physical points as positions of individuals and states as movement directions. Given $\delta(P_i(t), P_j(t)) = 1$ if $P_i(t), P_j(t)$ form the same triangulation (note that $\delta(P_i(t), P_i(t)) = 1$), otherwise it is zero, we have a strategy function for LRA as follows.

$$f_{\text{LRA}}(P^t, S^{t-1}, i) = \begin{cases} S_w(t), & \text{if } i = L \\ \frac{\sum_j S_j(t-1) \cdot \delta(P_i(t), P_j(t))}{\sum_j \delta(P_i(t), P_j(t))}, & \text{otherwise} \end{cases} \tag{8.2}$$

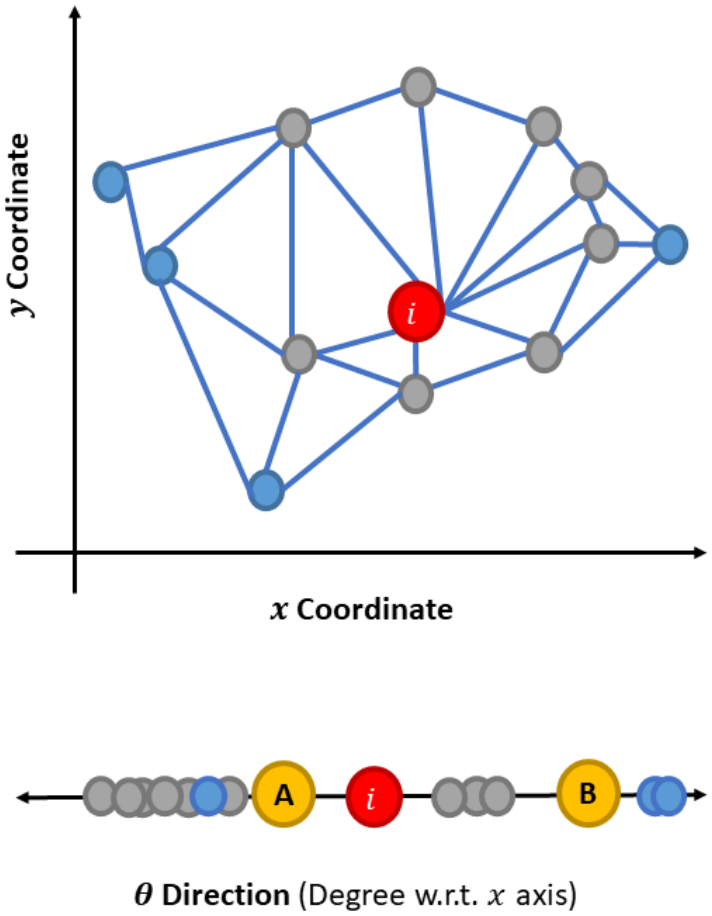


Figure 36. An example of physical points as positions and state points as directions. In position space (above), the individual i (red node) has all gray nodes as its neighbors in LRA since they are neighbors in Delaunay triangulation. In the direction space (below), i updates its next direction to be A rather than B since B is outside the i's neighbor convex hull.

The difference between f_{HM} (Equation 8.1) and f_{LRA} (Equation 8.2) is that f_{HM} infers the next state based on a *fixed* structure of a probabilistic following network $\mathcal{G}$, independently from the physical space P^t, whereas f_{LRA} predicts the next state based on the physical space P^t. In other words,, f_{LRA} represents an assumption that an agent follows a fixed set of specific individuals w.r.t. the preference graph $\mathcal{G}$ regardless of their relative physical position, while f_{LRA} represents an assumption that an agent follows anyone who happens to be around without any preference to follow specific individuals. The next theorem shows that the Local Reversible Agreement is ϵ-convergent.

Theorem 8.2.4 (Chazelle 2011 [47]) *For any $0 < \epsilon \leq \rho/n$, an n-agent reversible agreement system is ϵ-converged in time $\mathcal{O}(\frac{1}{\rho} \cdot n^2 log_2(\frac{1}{\epsilon}))$.*

According to the work by Chazelle [47], LRA is still converged even if one of the agents does not update. In our case, if $S_w(t)$ is the same for every time step, then the fixed agent is L who always has $S_L(t) = S_w(0)$.

Corollary 8.2.5 *the n-agent LRA that has $G(t)$ being created from Delaunay triangulation sets converges to a single point.*

Proof The graph $G(t)$ that is built from Delaunay triangulation is always connected. For each time step, each agent converges to the center of the neighbors' convex hull. Since everyone is connected and the system is ϵ-converge, by transitivity, the entire group converges to the single point.

In fact, if the fixed point is $S_w(0)$, then, at the equilibrium point, all states form a convex hull around $S_w(0)$ with the diameter at most ϵ [47]. In contrast, if $S_w(t)$ is not always the same, then the group moves following $S_w(t)$ with some time delay. The Corollary 8.2.5 tells us that if we follow our physical neighbors (e.g. directions) and everyone does the same thing, the entire group will reach the same consensus (moving to the same direction). In general, if $G(t)$ is strongly connected, everyone follows neighbors in $G(t)$, and there is one individual L who never follows anyone, then the group converges to L's state. Additionally, Corollary 8.2.5 is always true in any metric space where a Delaunay triangulation exists.

According to Corollary 8.2.5 and Proposition 8.2.1, if a target state never changes: $S_w(t) = S_w(0)$ for all $t > 0$, then the set of strategy functions $\mathcal{F}$ that contains only LRA strategy functions is a set of coordination strategies.

8.2.1.3 Discussion

According to Theorem 8.2.3, Corollary 8.2.5, Proposition 8.2.1, and Proposition 8.2.2, if the data has coordination behaviors, then either HM, LRA, or a mix of those strategies may be the cause of the coordination. However, the question still remains regarding how to infer which strategy is the cause of the coordination. In the next section, we propose a solution to address this question.

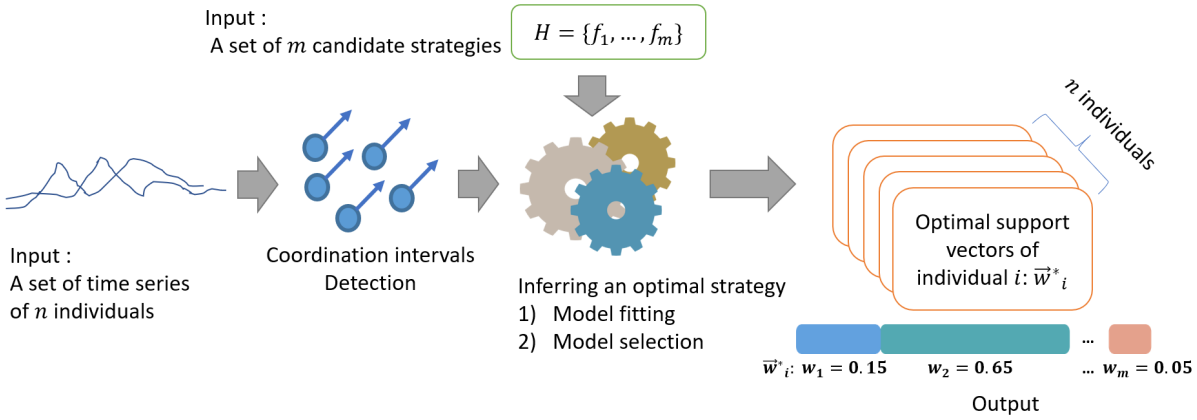


Figure 37. An overview of the proposed framework. Given a set of time series as inputs, first, the framework detects coordination intervals. Second, the framework infers the optimal strategy from a set of candidate strategies. In the model fitting step, the framework evaluates the performance of each pure and mixed strategy on the task of individual-level direction prediction using training datasets, then it gives the support weight to each strategy w.r.t. some minimum weight threshold $\vec{\kappa}$. The threshold $\vec{\kappa}$ represents the weight bias toward specific strategies. In the model selection step, the framework selects the best strategy that performed well in validation datasets from a set of various strategies that have been trained by different $\vec{\kappa}$. Finally, the framework reports the optimal support weight of each strategy for each individual. If the weight w_i of strategy f_i is high, then i might use f_i strategy to coordinate with the group. For simplicity of the exposition, in this paper, we deploy three candidate strategies: 1) an individual moves following specific individuals (HM), 2) an individual moves following its physical neighbors (LRA), and 3) an individual moves toward the same direction as in its previous time steps (Auto Regressive). However, other candidate strategies are also admissible.

8.3 Method

We are now ready to formally state our approach of inferring movement coordination strategies of agents represented by a collection of time series.

8.3.1 Setting

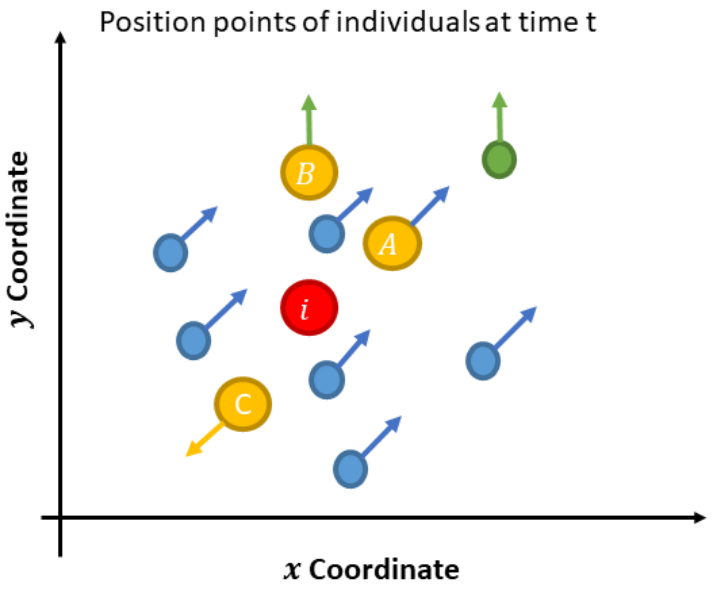


Figure 38. An example of movement strategy inference for i. Given the information on positions and directions of individuals in the past (blue and green nodes), we want to infer the i's strategy of movement that can be whether that i's next direction follows its neighbors (A node), or follows specific individuals (B node), or neither (C node).

We define a movement direction as a state, but the approach generalizes to arbitrary definitions of states. Hence, $\mathcal{S}_k$ is a set of time series of direction. We are assume the following are given as inputs: a set of possible strategy functions $\mathcal{H}$, a collection of position-time-series sets $\mathbb{P} = \{\mathcal{P}_k\}$ and a collection of direction-time-series sets $\mathbb{S} = \{\mathcal{S}_k\}$, where $\mathcal{P}_k = \{P_1, \ldots, P_n\}$ and $\mathcal{S}_k = \{S_1, \ldots, S_n\}$. A pair of $\mathcal{P}_k, \mathcal{S}_k$ represents kth coordination event that was generated by n agents moving in two-dimensional position space. Each $\mathcal{S}_k$ contains a coordination interval. The goal to to infer the set of strategy functions $\mathcal{F} \subseteq \mathcal{H}$ that generated $\mathcal{S}_k$. The overview of our framework is in Figure 37.

We use direction, rather than position, to define the state of an individual and the proxy for collective coordination. The main reason is that directional coordination is common in biology. For example, in [53], the authors report that a fish tends to imitate the direction of neighbors ahead to form collective movement, and other examples abound. Secondly, synchronization to the same direction implies a collective movement while synchronization to the same position implies staying in the same position without movement. In this paper, we focus on coordination of movement, therefore, we cannot use positions as states to infer strategies of movement. The final reason for defining states as directions is to use a dimension independent of the positions, which we use to define states of individual strategies. We need to differentiate between the strategy that an individual follows specific individuals' direction regardless of their physical neighbors' choices of direction versus the strategy that an individual follows their physical neighbors' direction without any preference to follow specific individuals.

For simplicity of the exposition we deploy three strategy functions for our framework: HM, LRA, and Auto regressive model (AR). Again, other candidate strategies are admissible. However, these three strategies are canonical exemplars since they allow to differentiate whether the strategy functions that generated a time series of directions of each agent is more hierarchical (HM), or it is more dependent on the physically proximity neighbors (LRA), or it is just a simple function of the agent's past history, independent of its neighbors. The example of movement strategy inference is in Figure 38.

We separate $\mathbb{P}$ and $\mathbb{S}$ to be a training part, ($\mathbb{P}_{\text{train}} \subset \mathbb{P}$, $\mathbb{S}_{\text{train}} \subset \mathbb{S}$), to perform a model fitting, and a validation part, ($\mathbb{P}_{\text{val}} = \mathbb{P} - \mathbb{P}_{\text{train}}$, $\mathbb{S}_{\text{val}} = \mathbb{S} - \mathbb{S}_{\text{train}}$) , to perform a model selection. In the case that the input is only a single physical time series $\mathcal{P}$, we use FLICA framework [1] to find coordination events and treat each event as a single $\mathcal{P}_k$. Hence, we have $\mathbb{P}$ containing multiple coordination events from $\mathcal{P}$. Then, we create a set of direction-time-series sets $\mathbb{S}$ from $\mathbb{P}$.

8.3.2 Model fitting

We concatenate all time series in $\mathbb{P}_{\text{train}}$ to be a single time series $\mathcal{P}_{\text{train}}$ and also concatenate $\mathbb{S}_{\text{train}}$ to be $\mathcal{S}_{\text{train}}$. Then we use $\mathcal{P}_{\text{train}}$, $\mathcal{S}_{\text{train}}$ to perform model fitting.

Before proceeding with the model fitting, the HM strategy function requires a probabilistic following network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. We infer $\mathcal{G}$ from $\mathcal{S}$ by using FLICA [1] to create a dynamic following network of $\mathcal{S}_{\text{train}}$. In this paper, the time window threshold of FLICA has been set at $\omega = 60$ time steps. In the next step, we find a global-leadership ranking, then we aggregate and normalize this dynamic network to be a DAG probabilistic network, such that the high-rank

agents do not have a probabilistic following edge to low-rank agents in $\mathcal{G}$. After we have $\mathcal{G}$, we calculate f'_{HM} as follows:

$$f'_{\text{HM}}(S^{t-1}, i) = \frac{S_i(t-1) + \sum_{(i,k)\in\mathcal{E}} p_{i,k} \cdot S_k(t-1)}{1 + \sum_{(i,k)\in\mathcal{E}} p_{i,k}}, \tag{8.3}$$

where $p_{i,k} \in [0, 1]$ is a probabilistic weight of edge $(i, k) \in \mathcal{E}$.

For the LRA strategy function, we use the same function as in Equation 8.2. Lastly, we apply auto regressive model to fit on $\mathcal{S}$ to represent f_{AR}. The f_{AR} predicts the next state of agent w.r.t. the average of the states from previous t' steps in $\mathcal{S}$. In this paper, we set $t' = 5$.

As mentioned before, we focus on three strategy functions: f_{HM} (Equation 8.3), f_{LRA} (Equation 8.2), and f_{AR}. We can view them as a mixed strategy, given a weight vector $\vec{w} = [w_1, w_2, w_3]^{\text{T}}$.

$$f_{\text{mix}}(a, \vec{w}) = w_1 f'_{\text{HM}}(a) + w_2 f_{\text{LRA}}(a) + w_3 f_{\text{AR}}(a) \tag{8.4}$$

Here $a = (P^t, S^{t-1}, i)$, w_1 is a support of HM, w_2 is a support of LRA, w_3 is a support of a linear regression model, and $w_1, w_2, w_3 \in [0, 1]$.

We us the sum square error (SSE) as our loss function. Our main goal is to find $\vec{w}^*$ that minimizes $risk(\mathcal{P}, \mathcal{S}, \vec{w}^*, i)$ below:

$$risk(\mathcal{P}, \mathcal{S}, \vec{w}, i) = \sum_{t=1}^{T} (D(t))^{\text{T}} \cdot (D(t)), \tag{8.5}$$

where $D(t) = |f_{\text{mix}}(P^t, S^{t-1}, i, \vec{w}) - S_i(t)|$ is a difference between predicted and actual direction in which agent i moved at time t.

For each agent i, given $\mathcal{P}_{\text{train}}, \mathcal{S}_{\text{train}}$ and a threshold vector $\vec{\kappa} = [k_1, k_2, k_3]^{\text{T}}$, we can find the optimal support vector $\vec{w}^*$ as the following optimization problem:

$$\begin{aligned} \underset{\vec{w}}{\text{minimize}} \quad & risk(\mathcal{P}_{\text{train}}, \mathcal{S}_{\text{train}}, \vec{w}, i) \\ \text{subject to} \quad & w_i \geq \kappa_i,\ w_i \in \vec{w}, \kappa_i \in \vec{\kappa}. \\ & \sum_i w_i = 1 \\ & w_i, \kappa_i \in [0, 1]. \end{aligned} \tag{8.6}$$

We use the Interior point algorithm [79], which is a large-scale algorithm, to solve the optimization Problem in Equation 8.6, which can be consider as a constrained linear least-squares problem. A threshold $\vec{\kappa}$ represents a model bias toward specific strategies. For example, if we have prior information that, with high probability, an agent i uses LRA strategy function, then we can set $\kappa_2 = 0.5$ to enforce the optimizer to vary the support w_2 within $[0.5, 1]$ interval instead of the $[0, 1]$ interval. The benefit of having $\vec{\kappa}$ is to prevent overfitting. For any agent i, suppose $\vec{w}^*_{i,k}$ is the optimal solution of the optimization problem in Equation 8.6 w.r.t. $\vec{\kappa}_k$, then we call $(\vec{w}^*_{i,k}, \vec{\kappa}_k)$ a model.

8.3.3 Model selection

First, we vary $\vec{\kappa}_k$ and find a model $(\vec{w}^*_{i,k}, \vec{\kappa}_k)$ for each agent i from $\mathcal{P}_{\text{train}}$, $\mathcal{S}_{\text{train}}$. As the result, we have a set of models $\Phi_i = \{(\vec{w}^*_{i,k}, \vec{\kappa}_{i,k})\}$ that is now used to perform model selection for an agent i.

We concatenate all time series in $\mathbb{P}_{\text{val}}$ to be a single time series $\mathcal{P}_{\text{val}}$ and also concatenate $\mathbb{S}_{\text{val}}$ to be $\mathcal{S}_{\text{val}}$. Finally, for each agent i, we find the optimal support vector $\vec{w}_i^*$ using the equation below:

$$\vec{w}_i^* = \underset{(\vec{w}_{i,k}^*, \vec{\kappa}_k) \in \Phi_i}{\operatorname{argmin}} \; risk(\mathcal{P}_{\text{val}}, \mathcal{S}_{\text{val}}, \vec{w}_k^*, i). \tag{8.7}$$

After we get the support vector $\vec{w}_i^* = [w_{i,1}^*, w_{i,2}^*, w_{i,3}^*]^{\text{T}}$, if $w_{i,1}^*$ is the highest support in $\vec{w}_i^*$, then we say that agent i uses the HM strategy function to coordinate with its group. If $w_{i,2}^*$ has the highest support, then we say that i follows its physical neighbors to coordinate with the group. If $w_{i,3}^*$ has the highest support, then i just follows its own linear path independently, and if i's path is the target path S_w then i is an informed agent. Lastly, if at least two of $w_{i,1}^*$, $w_{i,2}^*$, $w_{i,3}^*$ show significantly high weights, then we conclude that i uses a mixed strategy.

8.4 Experimental setup

We test our approach both on simulated adn on biological data.

8.4.1 Simulations

First we validate our Coordination Strategy Inference Problem framework on simulated datasets. We generated a set of time series of 2-dimensional positions $\mathcal{P}$ by four different sets of strategy functions: $\mathcal{H}_{\text{HM}}$,$\mathcal{H}_{\text{LRA}}$,$\mathcal{H}_{\text{HM\&LRA}}$, and $\mathcal{H}_{\text{MIX}}$. We define a set of state-time-series $\mathcal{S} = \{S_i\}$ as a set of time series of directional degrees of $\mathcal{P} = \{P_i\}$, where $P_i = (P_i(0), \ldots, P_i(T))$ is a time series of positions of an agent i; $S_i = (S_i(0), \ldots, S_i(T))$ is time series of directional degrees of an agent i derived from a position time series P_i; $S_i(t) \in (-180, 180]$ is a degree

angle between a direction vector $\vec{v}_i(t) = P_i(t) - P_i(t-1)$ and x-axis direction vector $[1, 0]^{\mathrm{T}}$. Note that we need to be careful also of the distance between any $S_i(t)$ and $S_j(t)$ since -179° and 180° has a degree difference of only 1° degree but certainly have different implications for coordination.

$$\mathrm{DIST}_{\mathrm{dir}}(S_i(t), S_j(t)) = \begin{cases} |S_i(t) - S_j(t)|, & \text{if } |S_i(t) - S_j(t)| \leq 180 \\ 360 - |S_i(t) - S_j(t)|, & \text{otherwise} \end{cases} \tag{8.8}$$

Where $\mathrm{DIST}_{\mathrm{dir}}(S_i(t), S_j(t)) \in [0, 180]$.

We have only $\mathcal{P}$ as an input for our framework since we can create $\mathcal{S}$ from $\mathcal{P}$. In all following simulated datasets, there are 20 agents and ID(1) is the informed agent. ID(1) creates the target path by uniformly and randomly choosing a fixed direction $S_w(0)$ as the initial state, then continuing to move in the direction of $S_w(0)$ until the end of coordination.

8.4.1.1 Hierarchical Model Dynamic System

In this system, we used a set of strategy function $\mathcal{H}_{\mathrm{HM}} = \{f_i\}$ to generate $\mathcal{P}_{\mathrm{HM}}$ where all f_i is f_{HM} (Equation 8.1). The parameter in this model is the following probability $\rho \in [0, 1]$. We set the probability weight of all edges in a probabilistic following network $\mathcal{G}$ equal to ρ. The communication network $G(t)$ generated by $\mathcal{G}$ is used to update the directional state $S_i(t)$ by the strategy function f'_{HM}. All 19 agents always follow only ID(1) with the probability ρ. In other words, all nodes have edges to ID(1) with the weight ρ in $\mathcal{G}$. For each coordination event, it lasts 400 time steps. So, $\mathcal{P}_{\mathrm{HM}} = \{P_1, \ldots, P_{20}\}$ s.t. $P_i = (P_i(0), \ldots, P_i(400))$. We vary

$\rho \in \{0.25, 0.50, 0.75, 1.00\}$. For each ρ, we generated 100 coordination events. In total, we have 400 datasets for this system.

8.4.1.2 Local Reversible agreement system

We created other 100 datasets of LRA system. We used a set of strategy function $\mathcal{H}_{\text{LRA}} = \{f_i\}$ to generate $\mathcal{P}_{\text{LRA}}$ where all f_i is f_{LRA} (Equation 8.2). For each dataset, it contains a set of time series of positions from 20 agents, $\mathcal{P}_{\text{LRA}} = \{P_1, \ldots, P_{20}\}$, where $P_i = (P_i(0), \ldots, P_i(400))$. All agents updates their state $S_i(t)$ corresponding to their local neighbors' states using a strategy function f_{LRA}.

8.4.1.3 Hierarchical and Local Reversible agreement system

We created other 100 datasets of HM & LRA coordination events by $\mathcal{H}_{\text{HM\&LRA}}$. We use this simulation to represents the group that has a coordination interval even if some agents use the HM strategy function but others use the LRA strategy function. For each dataset, it contains a set of position time series from 20 agents, $\mathcal{P}_{\text{HM \& LRA}} = \{P_1, \ldots, P_{20}\}$, where $P_i = (P_i(0), \ldots, P_i(400))$. The ID(1) is the informed agent. Agents who possess ID(2-10) use f_{HM} with $\rho = 1.00$. The rest of ID(11-20) agents uses f_{LRA}.

8.4.1.4 Mixed strategy system

Lastly, we created other 100 datasets of mixed strategy of coordination events. For each dataset, it contains a set of 20-agent position time series $\mathcal{P}_{\text{MIX}} = \{P_1, \ldots, P_{20}\}$ where $P_i = (P_i(0), \ldots, P_i(400))$ is time series of positions of agent i. The ID(1) is the informed agent. Other agents updates their state $S_i(t)$ corresponding to either f_{HM} with probability 0.5 or f_{LRA} with probability 0.5.

8.4.1.5 Evaluation

For each model, we performed 10-fold cross validation to evaluate the performance. For each round of cross validation, we have 100 datasets that can be separated into 45 training datasets, 45 validation datasets, and 10 testing datasets. We concatenated all time series in $\mathbb{P}_{\text{test}}$ to be a single time series $\mathcal{P}_{\text{test}}$ and also concatenate $\mathbb{S}_{\text{test}}$ to be $\mathcal{S}_{\text{test}}$. Then we use $\mathcal{P}_{\text{test}}, \mathcal{S}_{\text{test}}$ to evaluate the direction prediction performance. We compare four strategy functions: f_{HM}, f_{LRA}, f_{AR}, and f_{OPT}, which is our framework optimal strategy function derived from Equation 8.4 and Equation 8.7. We use the risk function that has Equation 8.8 as a loss function to evaluate the model performance.

$$risk(\mathcal{P}, \mathcal{S}, f, i) = \frac{1}{T} \sum_{t=1}^{T} \text{DIST}_{\text{dir}}(S_i(t), f(S^{t-1}, P^t, i)) \tag{8.9}$$

For each agent i, the best fitting model is the model that minimizes the risk function $risk(\mathcal{P}_{\text{test}}, \mathcal{S}_{\text{test}}, f, i)$ in Equation 8.9.

$$f_i^* = \underset{f \in \{f_{\text{HM}}, f_{\text{LRA}}, f_{\text{AR}}, f_{\text{OPT}}\}}{\text{argmin}} risk(\mathcal{P}_{\text{test}}, \mathcal{S}_{\text{test}}, f, i) \tag{8.10}$$

For each strategy function f, we report the distribution of loss values of direction prediction from all agents in each time step as well as the group's average optimal support $\vec{w}_i^*$ from Equation 8.7. If the framework performs well, then it should give the highest support for the model that was used to generate the dataset.

8.4.2 Baboon behavioral experiment

The dataset is the recording of GPS collars of olive baboons (*Papio anubis*) troop in the wild in Mpala Research Centre, Kenya [6, 7] (see Section 4.2.1). We extracted coordination events by FLICA varying the network density threshold at 25th, 50th, 75th, and 99th percentile and the time window at 240 time steps to infer coordination events and 60 time steps to infer a dynamic following network. We used the 10-fold cross validation to report the results. For each round of cross validation, it has 45% of training, 45% of validation, and 10% of testing coordination events. For the rest of the evaluation steps, we evaluated and report the results the same way as described in **Evaluation Section**.

8.4.3 Fish behavioral experiment

We used the time series of golden shiners (*Notemigonus crysoleucas*) fish positions from [46] (see Section 4.2.2). The dataset was initially created to study information propagation via the fish visual fields [46].

In total, there were 24 trails of fish position time series $\mathbb{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_{24}\}$ in 2-dimensional space. For each $\mathcal{P}_k$, it consists of 70 fish, with 10 trained fish who are considered to be informed agents in our setting. On average, the time series in $\mathcal{P}_k$ has its length around 600 time steps. The trained fish moved toward the feeding site (the target path) and the group follows them. Due to the lack of information of identity for each individual in the different trails, we cannot train our framework in this dataset. Hence, we use fish data to demonstrate how to apply our framework to compare performance of each candidate strategy on direction prediction.

We compared the Informed strategy function f_{TF} against f_{LRA} in Equation 8.2. For each time step, f_{TF} updates $S_i(t)$ for any agent i from the average of $S_j(t)$ where j is a trained fish. We use the risk function in Equation 8.9 to compare the performance among these strategy functions. For each strategy function f, we report the distribution of all agents' direction prediction error in each time step from $\text{DIST}_{\text{dir}}(S_i(t), f(S^{t-1}, P^t, i)$.

8.4.4 Comparison with the state of the art method

Our method is the first approach to infer individual-level strategies that lead to group-level coordination. Thus, we compare our framework with the-state-of-the-art method, FLICA [3], for the task of leadership model classification. Since FLICA cannot infer the individual-level strategy, we evaluate both frameworks at the group-level classification task. We use simulated datasets from Section 8.4.1. Each set of time series has its label from one of the four models: HM, LRA, HM & LRA, and Mix strategy model. FLICA maps each set of time series to the leadership ranking and convex hull features. In our framework, we use the median of $\vec{w}_i^*$ (Equation 8.7) to represent the feature vector of each dataset. We use 10-fold cross validation on Random Forests [74] to report the evaluation results for both frameworks.

8.5 Results

We now report the results of our experiments.

8.5.1 Simulations

The results of inferring the coordination strategy in simulated datasets are shown in Table XXII. A row represents the results from datasets generated by a specific model. A column represents a strategy prediction error measured in degree units $[0°, 180°]$. OPT is the optimal

TABLE XXII

The result of predicting the direction of movement via 10-fold cross validation. We compared the result of our framework (OPT) against the base-line pure strategies: HM, LRA, and AR (auto regressive strategy). (*indicates the STD $\geq 20°$)

	Average degree prediction error $[0°, 180°]$			
Datasets\Strategies	OPT	HM	LRA	AR
HM	**12.40**	**12.98**	20.49	30.21*
LRA	**7.77***	16.93*	**7.76***	13.78*
HM & LRA	**4.42**	13.39*	13.59	23.87*
Mixed Str.	**29.33***	30.53*	31.69*	46.28*
Random	**89.74***	90.11*	89.70*	90.21*
Baboon	**53.16***	**53.16***	72.36*	85.84*

strategy function trained by our framework. HM is Equation 8.3. LRA is Equation 8.2. AR is the auto regressive strategy function that chooses the current direction t based on the previous five time steps from the same agent. We use AR as the baseline. In all datasets, our framework (OPT column) has the smallest error among all other strategies. For the first two rows of HM and LRA datasets, OPT has almost the same performance as the strategies used to generate

the data (HM row/column and LRA row/column). For HM & LRA datasets in the third row, each individual might use either HM or LRA strategy. Hence, using the homogeneous strategy to predict directions for all agents results in larger error (HM and LRA column). On the contrary, our framework can detect which individual uses which strategy. Hence, OPT performed better than all pure strategies. Similarly, for the mixed strategy datasets (Mixed Str. row), each individual might use either HM or LRA as its strategy with the probability 0.5. Since our framework can infer mixed strategies, it performed better than using any pure strategy. Lastly, we reported the results of the direction prediction from the 100 datasets of time series generated from n agents moving uniformly and randomly in any direction (Random row). The result shows that all strategies included in our framework produced the same bad result with the loss value at 90° degree. This shows that our framework does not find an artifact model where none exists.

Table XXIII shows the support vectors for each strategy corresponding to the datasets in Table XXII in the OPT column. For each element in the table, the first number is the predicted support from our framework and the second is the actual support that we used to create the datasets. For example, in the first element of HM row, 0.85/1.00 means we used HM strategy to create HM datasets and the framework inferred the HM support in these datasets as 0.85. Overall, our framework correctly inferred the support vectors of all non-random datasets, while avoiding overfitting.

TABLE XXIII

The average optimal support vector $\vec{w}$ of all agents from 10-fold cross validation, inferred by our framework from simulated and the Baboon datasets.

	Average Support $\vec{w}$ (predict/actual)		
Datasets	w_1:HM	w_2:LRA	w_3:AR
HM	**0.85/1.00**	0.12/0.00	0.03/0.00
LRA	0.02/0.00	**0.98/1.00**	0.00/0.00
HM & LRA (HM part)	**1.00/1.00**	0.00/0.00	0.00/0.00
HM & LRA (LRA part)	0.00/0.00	**1.00/1.00**	0.00/0.00
Mixed Strategy	**0.48/0.50**	**0.48/0.50**	0.04/0.00
Random	0.09/0.00	0.86/0.00	0.05/0.00
Baboon	1.00/NA	0.00/NA	0.00/NA

8.5.2 Baboon behavioral experiment

We varied the threshold of the following network density to infer coordination events in the baboon dataset. We report the average result from all the thresholds. The last row of

Table XXII shows the result of the direction prediction of baboons, using different coordination strategies. The OPT coordination strategy, as derived by our framework, is in the last row in Table XXIII. According to the result, OPT used HM as the pure strategy. The errors of HM and OPT strategies indicate that baboons have a strong preference to follow a pre-determined individual or a set of individuals, rather than their neighbors in the position space. This is consistent with the biological understanding of the baboon social behavior.

8.5.3 Fish behavioral experiment

TABLE XXIV

Comparison between LRA and Informed strategies to predict movement directions of 24 trails of fish.

	Error of degree prediction $[0°, 180°]$	
Strategies	Mean	STD
LRA	41.51	45.11
Informed Strategy	54.46	47.68

The results of the direction prediction in fish dataset, for LRA and Informed strategies, are in Table XXIV. The LRA performed better than the Informed strategy, indicating that fish

TABLE XXV

The results of model classification of FLICA and the proposed framework via 10-fold cross validation. We use Random Forest as the main classifier for both frameworks.

	FLICA			**Proposed framework**		
Classes	**Prec.**	**Rec.**	**F1**	**Prec.**	**Rec.**	**F1**
HM	1	0.75	0.86	1	1	1
LRA	0.8	1	0.89	1	1	1
HM & LRA	0.94	1	0.97	0.98	1	0.99
Mixed Str.	0.90	0.94	0.92	1	0.98	0.99
Random	1	0.9	0.95	1	1	1

follow their immediate neighbors in space. This result is supported by the work in [46], showing that fish do not directly know who leads the group but follow their neighbors.

8.5.4 Comparison with the state of the art method

The result of model classification using FLICA as well as the proposed framework is in Table Table XXV. In all datasets, the proposed framework performed better than FLICA. This indicates that the group-level features that FLICA provides for classification are not sufficiently informative to be used to categorize complicated datasets where individuals may use a heterogeneous set of strategies (e.g. HM & LRA).

8.6 Conclusions

In this work, we formalized a new computational problem, Coordination Strategy Inference Problem. Given a set of candidate strategies and a set of time series of coordinated movement as inputs, our goal is to infer the original strategy that each individual used to achieve the group coordination. We showed that a strategy that has the convergence property can guarantee that the group reaches coordination. We provide the first methodology (to the best of our knowledge) to infer the set of strategies that each individual uses to achieve movement coordination at the group level. We evaluated and demonstrated our framework performance in simulated datasets as well as two biological datasets: baboon and fish. Our framework was able to infer the original set of strategy functions that generated each simulated dataset. The results show that our approach is highly accurate in inferring the correct strategy in simulated datasets even in complicated mixed strategy settings. Moreover, our framework performed classification of group-level coordination models from time series better than FLICA framework, which is the-state-of-the-art approach for the task. Animal data experiments show that fishes, unsurprisingly, follow their neighbors, while baboons have a preference to follow specific individuals. Although we used the specific setting of focusing on the direction of movement as the definition of an agent's state and used three exemplar candidate strategy, our methodology easily generalizes to arbitrary time series data, beyond movement data, and other candidate strategies. While for the fairness of comparison with the biological datasets we used simulated data of 20 individuals, it is clear that there are no inherent limitations in the approach to scale to much larger datasets. The only barrier is the availability of data. The code and datasets

that we used in this work can be found at [80]. In the next chapter, we will introduce the proposed framework, which is the extension of mFLICA framework in Chapter 6. The proposed framework can be used to infer leadership dynamics from time series data.

CHAPTER 9

MINING AND MODELING COMPLEX LEADERSHIP DYNAMICS OF MOVEMENT DATA

9.1 Introduction

> MINING PATTERNS OF LEADERSHIP DYNAMICS: Given time series of individual activities, the goal is to mine and model frequent patterns of leadership dynamics, including emergence of multiple leaders, convergence of multiple leaders to a single one, or change of a leader.

Leadership is an essential part of collective decision and organization in social animals, including humans. In nature, leadership is dynamic and varies with context or temporal factors. Understanding dynamics of leadership, such as how leaders change, emerge, or converge, allows scientists to gain more insight into group decision-making and collective behavior in general. However, given only data of individual activities, it is challenging to infer these dynamic leadership events. In this work, we focus on mining and modeling frequent patterns of leadership dynamics. In Chapter 3, we formalized a new computational problem, MINING PATTERNS OF LEADERSHIP DYNAMICS. In this chapter, we propose a framework as a solution of this problem. Our framework can be used to address several questions regarding leadership dynamics of group movement. We use the leadership inference framework, mFLICA, to infer the time series

of leaders from movement datasets, then propose the approach to mine and model frequent patterns of leadership dynamics.

We evaluate our framework performance by using several simulated datasets, as well as using the real-world dataset of baboon movement to demonstrate the application of our framework. There are no existing methods to address this problem, thus, we modify and extend the existing leadership inference framework to provide a non-trivial baseline. Our framework performs better than this baseline in all datasets. Moreover, we also propose a method to perform statistical significance tests, comparing inferred frequent patterns of leadership dynamics with our proposed null hypotheses. Our framework opens the opportunities for scientists to generate scientific hypotheses that can be tested statistically regarding dynamics of leadership in movement data.

9.2 Methods

To solve Problem 4, Mining Patterns of Leadership Dynamics, we propose the framework consisting of four parts (Figure 39.) Given a set of time series of movement $\mathcal{U} = \{U_1, \ldots, U_n\}$, where $U_i \in \mathcal{U}$ is a two-dimensional time series of length T, first, we infer a dynamic following network and time series of leaders $\mathcal{L}$ using mFLICA framework [4] (Chapter 6 or Section 2.2). Second, we infer a diagram of leadership dynamics $\mathcal{T}$ from $\mathcal{L}$ in Section 9.2.2. Third, we detect the sequence patterns on $\mathcal{L}$ in Section 9.2.3. Finally, we deploy hypothesis tests to evaluate significance of leadership dynamics compared to our proposed null models in Section 9.2.4.

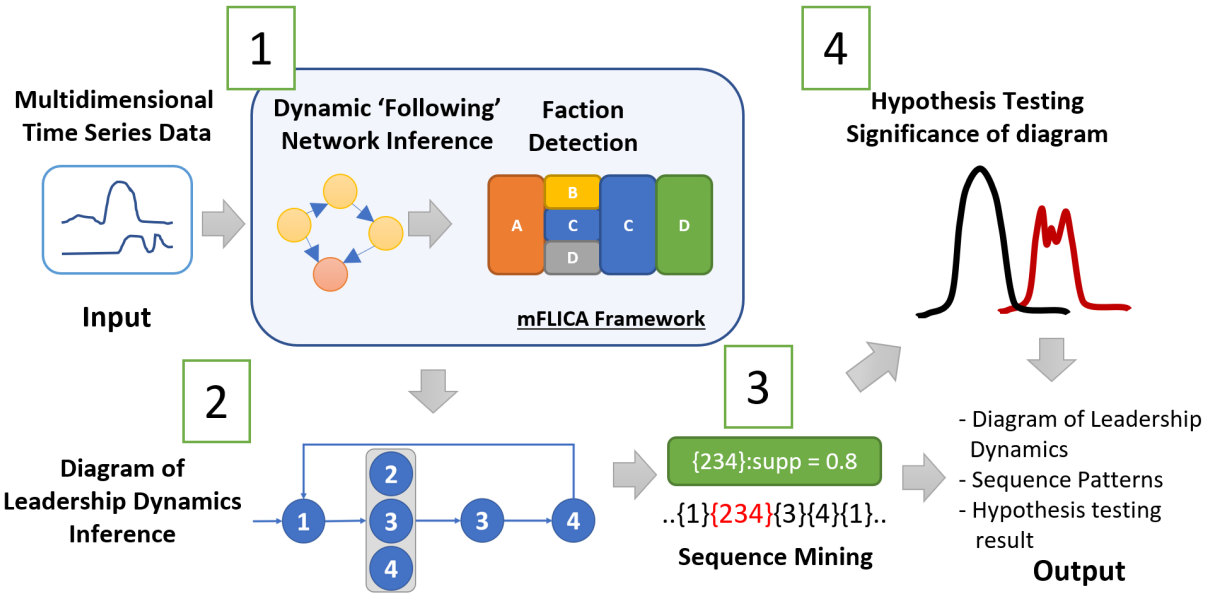


Figure 39. A high-level overview of the proposed framework.

9.2.1 Recap of mFLICA

In this section, we provide the brief details of mFLICA (see Chapter 6 for more details) as a reminder. Given a pair of time series U and Q, mFLICA uses Dynamic Time Warping (DTW) [76] to infer a following relation. Suppose $P_{U,Q}$ is an optimal warping path from DTW dynamic programming matrix, where $(i, j) \in P_{U,Q}$ implies $U(i)$ matched with $Q(j)$ in the matrix. Intuitively, if U is followed by Q with the time delay $\Delta_{i,j}$, then $j - i = \Delta_{i,j}$. Hence, we can compute the *following* relation by the equation below.

$$\mathrm{f}(P_{U,Q}) = \frac{\sum_{(i,j)\in P_{U,Q}} \mathrm{sign}(j-i)}{|P_{U,Q}|}. \tag{9.1}$$

Suppose we have a similarity threshold σ, there we say there is a following relation if $|\mathrm{f}(P_{U,Q})| \geq \sigma$, where Q follows U if $\mathrm{f}(P_{U,Q}) \geq \sigma$ and U follows Q if $\mathrm{f}(P_{U,Q}) \leq -\sigma$. We set $\sigma = 0.5$ as a default.

Next, given a time window ω and a sliding window parameter $\delta = 0.1\omega$, we have the i^{th} time window interval $w(i) = [(i-1) \times \delta, (i-1) \times \delta + \omega]$. mFLICA creates a following network for each set of time series within interval $w(i)$ of $\mathcal{U}$. An edge of a following network is inferred by Equation 9.1 with the weight $|\mathrm{f}(P_{U,Q})|$. Hence, after every interval $w(i)$ has its following network, we have a dynamic following network $\mathcal{G} = \langle G_t \rangle$ of $\mathcal{U}$.

Lastly, for each time step t, mFLICA uses Breadth First Search (BFS) to infer factions and initiators within a following network G_t. The faction initiators are nodes with out-degree zero and in-degree non-zero. By applying BFS to dynamic following network $\mathcal{G}$, we have the time series of leaders $\mathcal{L} = (\mathcal{L}(1), \ldots, \mathcal{L}(T))$ as the output of this step.

9.2.2 Inferring transition diagram of leadership dynamics

We use Hidden Markov Model (HMM) [64] to model a diagram of leadership dynamics $\mathcal{T} = (V_{\mathcal{T}}, E_{\mathcal{T}})$ in Def. 17 and use Baum–Welch algorithm [81] to infer the maximum likelihood estimates of parameters of HMM from the time series of leader $\mathcal{L}$. In this setting, we have a set of frequent-leader sets $\mathcal{S}_{\mathcal{L}}$ as a set of states in HMM with the support threshold $\phi = 0.01$. In HMM, the stochastic transition matrix A, which has its size $|\mathcal{S}_{\mathcal{L}}| \times |\mathcal{S}_{\mathcal{L}}|$, describes estimated probabilities that a group changes its current set of leaders to another set of leaders (e.g. group merging or splitting.) However, since we are interested only in the events of state changes, we

ignore the self-transition probability and normalize A to be A^* (Equation 9.2), which is the adjacency matrix of $\mathcal{T}$.

Given a time series of leaders $\mathcal{L}$, we can easily infer a set of leader sets $\mathcal{S}_\mathcal{L}$. Then, let $\mathcal{S}_{\text{HMM}}$ be a set of states in HMM where $\mathcal{S}_{\text{HMM}}$ and $\mathcal{S}_\mathcal{L}$ are in one-to-one correspondence. We represent each state in S_{HMM} as a number in $[1, |\mathcal{S}_{\text{HMM}}|]$, then we create $\mathcal{L}_{\text{HMM}}$ by replacing each element in $\mathcal{L}$ with the number of corresponding state in $\mathcal{S}_{\text{HMM}}$. For example, in Figure 40 (Dynamic Type 1), we have a set of leader sets $\mathcal{S}_\mathcal{L}$ and $\mathcal{S}_{\text{HMM}}$, where $\{\text{ID1}\}, \{\text{ID2}, \text{ID3}, \text{ID4}\}, \{\text{ID3}\}$, and $\{\text{ID4}\}$, in $\mathcal{S}_\mathcal{L}$ have corresponding elements in $\mathcal{S}_{\text{HMM}}$ as $1, 2, 3$ and 4, respectively.

Initially, we set a stochastic transition matrix $A = \{a_{i,j}\}$ (i, j are the states) and the initial state distribution π_i uniformly. We have the set of observation values $Y = \{1, \ldots, |\mathcal{S}_{\text{HMM}}|\}$. In this setting, there is no hidden state since an observation value is an identity of a state. However, in HMM, at any state i, there is a required probability $b_{i,j}$ of observing value j at the state i (typically represented by a matrix $B = \{b_{i,j}\}$.) Here, the probability $b_{i,j} = 1$ if $i = j$ and zero otherwise.

We use Baum–Welch algorithm [81] to infer $A = \{a_{i,j}\}$, then we normalize A to create $A^* = \{a^*_{i,j}\}$ by the equation below.

$$a^*_{i,j} = \begin{cases} 0, & i = j \\ \frac{a_{i,j}}{\sum_{k=1,k\neq j}^{|S_{\text{HMM}}|} a_{i,k}}, & \text{Otherwise.} \end{cases} \tag{9.2}$$

9.2.3 Mining sequence patterns of leadership dynamics

After having a diagram of leadership dynamics $\mathcal{T} = (V_{\mathcal{T}}, E_{\mathcal{T}})$, for each pair of nodes $(i, j) \in V_{\mathcal{T}}$, we find a sequence pattern, which is a path $P_{i,j} = (v(1) = i, \ldots, v(k) = j)$, where for all $u \in V_{\mathcal{T}}$, $a^*_{v(t-1),v(t)} > a^*_{v(t-1),u}$.

$P_{i,j}$ is an order sequence that the previous state $v(t-1) \in P_{i,j}$ has the highest probability to change to the next consecutive state $v(t) \in P_{i,j}$, given a starting point at i and the final state at j.

Given $A^* = \{a^*_{i,j}\}$ as an adjacency matrix of $\mathcal{T}$, we convert A^* to be $A' = \{a'_{i,j}\}$ where $a'_{i,j} = \frac{1}{a^*_{i,j}}$. Then, we use the standard Dijkstra's algorithm to find the shortest path between every two nodes in A'. Hence, $P_{i,j}$ is the shortest path between i and j in A'[1]. Let ν be a number of times that the full sequence of $P_{i,j}$ occurs in $\mathcal{L}$ and N be a number of times that leadership state change happens in $\mathcal{L}$ (*e.g.*, two sub-groups merged together, changing the leader), we can find the support of $P_{i,j}$ in the time series of leader $\mathcal{L}$ by the equation below:

$$\text{supp}_{path}(\mathcal{L}, P_{i,j}) = \frac{\nu \times (|P_{i,j}| - 1)}{N}. \tag{9.3}$$

Specifically, ν is a number of times that all pairs of nodes $v(t-1), v(t) \in P_{i,j}$ s.t. $v(t-1)$ appear before $v(t)$ in $P_{i,j}$ also appear in $\mathcal{L}$.

TABLE XXVI

Details of non-parametric tests used in this paper. A significant level has been set at $\alpha = 0.01$ for all experiments.

Method	**Null hypothesis H_0**
Kolmogorov-Smirnov test [82]	Two samples are from the same distribution
Wilcoxon Rank Sum Test [83]	
Kruskal-Wallis Test [84]	

9.2.4 Hypothesis testing

9.2.4.1 Evaluating the significance of leadership-event order

Given a time series of leaders $\mathcal{L}$ and a diagram of leadership dynamics $\mathcal{T}$ inferred from $\mathcal{L}$, we perform a random permutation of elements in $\mathcal{L}$ to create $\mathcal{L}_{\text{rand}}$, then we infer a diagram of leadership dynamics $\mathcal{T}_{\text{rand}}$ from $\mathcal{L}_{\text{rand}}$ by the method described by the previous section. Afterwards, we test the similarity of the edge-weight distributions of $\mathcal{T}$ and $\mathcal{T}_{\text{rand}}$. We deploy three non-parametric methods, shown in Table XXVI, to perform the tests. If all three methods successfully the corresponding reject the null hypothesis with $\alpha = 0.01$, then we conclude that

[1]Note, this can be done since the probability condition is independent of each pair and not cumulative over the path

the edge-weight distribution of $\mathcal{T}$ is significantly different from $\mathcal{T}_{\text{rand}}$'s although the support value of each node in both graphs are the same.

9.2.4.2 Evaluating the significance of frequencies of leadership-event sequences

After finding all the sequences for every pair of nodes in Section 9.2.3, we compute the support $\text{supp}_{path}(\mathcal{L}, P_{i,j})$ of each sequence $P_{i,j}$. This gives the sequence-support distribution of $\mathcal{T}$. Next, we rewire $\mathcal{T}$ to be $\mathcal{T}_{\text{rand}}$ by uniformly and randomly changing the end points of each edge in $\mathcal{T}$, then we calculate the sequence-support distribution of $\mathcal{T}_{\text{rand}}$ (Equation 9.3.) Lastly, we test whether $\mathcal{T}$ and $\mathcal{T}_{\text{rand}}$ sequence-support distributions are different the same way as in the previous section.

We repeat both types of significance tests 100 times and report the percentage of times that the tests successfully reject H_0 for each dataset.

9.2.5 Time and space complexity

The time complexity of mFLICA is $\mathcal{O}(n^2 \times \omega \times T)$, where n is a number of time series, T is a length of time series, and ω is a time window parameter. The time complexity of Baum−Welch algorithm to infer a diagram of leadership dynamics is $\mathcal{O}(m^2 \times T)$ where m is the number of frequent-leader sets. Typically, $m < n$ since there are fewer frequent-leader sets than individuals. Hence, our framework's overall time complexity is $\mathcal{O}(n^2 \times \omega \times T)$. For the space complexity, the part that requires space the most in our framework is the space for the dynamic following network that costs $\mathcal{O}(n^2 \times T)$.

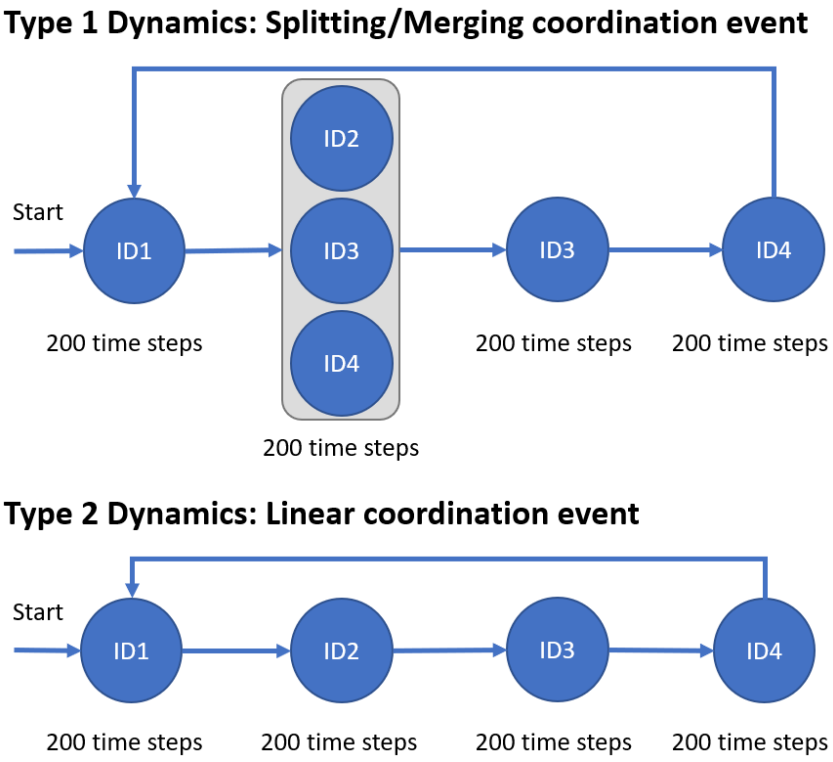


Figure 40. Splitting/Merging (above) and Linear (below) coordination event. Each node represents the ID of leader of each sub-group at the particular time and each edge represents the change of group's leaders.

9.3 Evaluation Datasets

We evaluate our method on synthetic datasets generated using a variety of leadership models with a variety of patterns of leadership dynamics. We directly use all datasets from mFLICA work in Chapter 6 (see Section 6.3.2). Nevertheless, we recap the types of coordination dynamics here below.

9.3.1 Type 1 Dynamics: Splitting/Merging coordination event

In this type of coordination event (Figure 40 above), ID1 leads the entire group for 200 time steps. Then, the group splits into three equal size sub-groups lead by ID2, ID3, and ID4, for the duration of 200 time steps. Afterwards, all sub-groups are merged into a single group again lead by ID3 for another 200 time steps. Finally, ID4 leads the entire group for the last 200 time steps.

9.3.2 Type 2 Dynamics: Linear coordination event

In this type of coordination event (Figure 40 below), ID1 leads first, then ID2 leads, ID3 leads after ID2, and ID4 leads after ID3. Each leader leads the group for 200 time steps.

9.4 Evaluation criteria

In simulated datasets, we compare the inferred adjacency matrix $A = \{a_{i,j}\}$ of a digraph of leadership dynamics $\mathcal{T} = (V_{\mathcal{T}}, E_{\mathcal{T}})$ against the ground truth matrix $A^* = \{a^*_{i,j}\}$. For the Splitting/Merging coordination event, the ground-truth set of frequent-leader sets is $\mathcal{S}^*_{\mathcal{L}} = \{\{\text{ID1}\}, \{\text{ID2}, \text{ID3}, \text{ID4}\}, \{\text{ID3}\}, \{\text{ID4}\}\}$. All elements in A^* are zero except $a^*_{\{\text{ID1}\},\{\text{ID2,ID3,ID4}\}} = 1, a^*_{\{\text{ID2,ID3,ID4}\},\{\text{ID3}\}} = 1, a^*_{\{\text{ID3}\},\{\text{ID4}\}} = 1$, and $a^*_{\{\text{ID4}\},\{\text{ID1}\}} = 1$. For the Linear coordination event, $\mathcal{S}^*_{\mathcal{L}} = \{\{\text{ID1}\}, \{\text{ID2}\}, \{\text{ID3}\}, \{\text{ID4}\}\}$ and all elements in A^* are zero except $a^*_{\{\text{ID1}\},\{\text{ID2}\}} = 1, a^*_{\{\text{ID2}\},\{\text{ID3}\}} = 1, a^*_{\{\text{ID3}\},\{\text{ID4}\}} = 1$, and $a^*_{\{\text{ID4}\},\{\text{ID1}\}} = 1$.

Let $\mathcal{S}_{\mathcal{L}}$ and $\mathcal{S}^*_{\mathcal{L}}$ be the predicted and the ground truth sets of frequent-leader sets, respectively. The loss function of A and A^* is below:

$$loss(A, A^*) = \frac{\sum_{i,j \in \mathcal{S}^*_{\mathcal{L}} \cap \mathcal{S}_{\mathcal{L}}} |a_{i,j} - a^*_{i,j}| + \mathrm{FP}(A, A^*) + \mathrm{FN}(A, A^*)}{n_{A^*}} \tag{9.4}$$

$$\mathrm{FP}(A, A^*) = \sum_{i,j \in \mathcal{S}_{\mathcal{L}} \setminus \mathcal{S}^*_{\mathcal{L}}} |a_{i,j}| \tag{9.5}$$

$$\mathrm{FN}(A, A^*) = \sum_{i,j \in \mathcal{S}^*_{\mathcal{L}} \setminus \mathcal{S}_{\mathcal{L}}} |a^*_{i,j}| \tag{9.6}$$

Where n_{A^*} is the number of elements within A^*. The first term in Equation 9.4 represents the $L1$-norm difference between each element in A and A^* (probabilities) when the predicted states are the same as the ground truth. The second term represents the false positive case when the framework predicts the states that do not exist in the ground truth. The last term represents the false negative case when the framework misses prediction of a state that exists in the ground truth.

9.5 Results

We set the time window parameter ω using the inference method in [4]. Figure 41 and Figure 42 show the examples of inferred diagrams of leadership dynamics by our framework from Type-1-HM (Hierarchical model with Splitting/Merging coordination events) and Type-2-HM (Hierarchical model with Linear coordination events) datasets respectively. In Figure 41, comparing the inferred diagram with the ground truth, only nodes $\{2, 4\}$ and $\{2, 3\}$ are false positive nodes.

This implies that despite the complex dynamics of leadership in Type-1-Dynamics case, our framework was still able to retrieve the diagram of leadership dynamics accurately. For the Type-2-HM dataset, which is less complex than Type1-HM case, Figure 42 shows that there are no false positive nodes in the inferred diagram. Moreover, in both Type-1-HM and Type-2-HM cases, the support of each node should be 0.25, and our framework can infer the support for each node closely to 0.25.

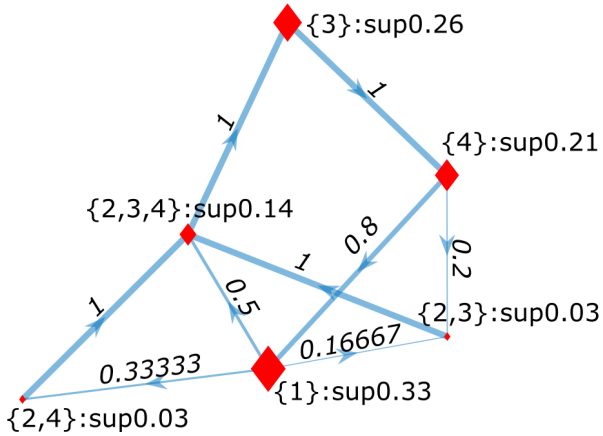


Figure 41. The example of the inferred diagram of leadership dynamics by our framework from a Type-1-HM dataset. Comparing the inferred diagram with the ground truth, only nodes $\{2, 4\}$ and $\{2, 3\}$ are false positive nodes. The support of $\{1\}, \{2, 3, 4\}, \{3\}$ and $\{4\}$ should be 0.25, and our framework can infer the support for each node closely to 0.25.

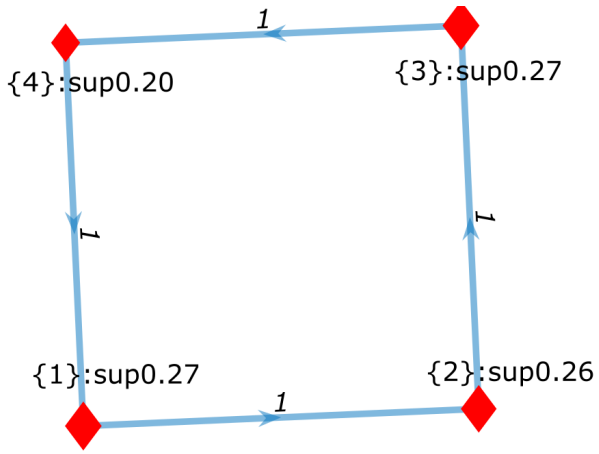


Figure 42. The example of the inferred diagram of leadership dynamics by our framework from a Type-2-HM dataset. Comparing the inferred diagram with the ground truth, there are no false positive nodes. The support of $\{1\}, \{2\}, \{3\}$ and $\{4\}$ should be 0.25, and our framework can infer the support for each node closely to 0.25

Regarding the mining sequence patterns of leadership dynamics described in Section 9.2.3, Table XXVII shows an example of max-support sequences of leadership dynamics that our framework reported from HM datasets. In both dynamics types, the sequences are consistent with the ground truth in Figure 40.

Next, we compared our framework, which uses the following networks concept [4], to the method based on direction networks proposed in FLOCK method [30] to infer a diagram of leadership dynamics. In direction networks, at any time t, if i is moving toward the same

TABLE XXVII

The example of sequences of leadership dynamics that have the highest support from HM datasets.

Datasets	**Sequences**	$\text{supp}_{path}(\mathcal{L}, P_{i,j})$
Type-1-HM	{2,3,4},{3},{4},{1}	0.71
Type-2-HM	{1},{2},{3},{4}	0.95

direction as j but j is at the front of i, then i follows j. The median of all loss distributions in both Type-1 and Type-2 dynamics datasets are reported in Table XXVIII. The first row of Table XXVIII shows the distribution of loss values (Equation 9.4) in type-1-dynamics datasets. The direction network approach was reasonably competitive for the type 1 dynamics. We were able to use the direction networks to infer the states with splits and merges but the change of leadership was often missed by this underlying method. Not surprisingly, then, the direction network-based method performed significantly worse than the following network-based approach for the type-2-dynamics. Qualitatively, and as a distribution of the loss values overall, the following networks as the basis for the diagram inference performed better than the direction networks in our framework. In the second row of Table XXVIII, the following networks also performed better than direction networks in Type-2-dynamics datasets.

TABLE XXVIII

The median of loss values in the prediction task of diagrams of leadership dynamics.

Dyn. Type	**Type 1**			**Type 2**		
Model	**HM**	**DM**	**IC**	**HM**	**DM**	**IC**
Following Network	0.13	0.19	0.24	0	0.03	0.08
Direction Network	0.19	0.19	0.25	0.19	0.19	0.25

In Table XXIX, we reported the hypothesis testing results of the significance of leadership-event order (Section 9.2.4.1). With respect to the type of the leadership model, for the HM, which is a well-structure model, the inferred diagrams are more significantly different from the null-model diagram than for the other leadership models. With respect to the types of the dynamics, in the complex type-1-dynamics datasets our framework inferred diagrams that are more significantly different from the null model. Lastly, the following networks were able to infer diagrams that are more different from the null model than the direction networks.

For hypothesis testing of the significance of frequencies of leadership-event sequences (Section 9.2.4.2), the result is shown in Table XXX. Similar to the the edge-weight distribution

TABLE XXIX

Hypothesis testing results of the significance of leadership-event order in Section 9.2.4.1. We reject H_0 at $\alpha = 0.01$. Each element in the table represents the percentage of the times when the tests successfully reject H_0.

Dyn. Type	**Type 1**			**Type 2**		
Model	**HM**	**DM**	**IC**	**HM**	**DM**	**IC**
Following Network	0.99	0.55	0.38	0.86	0.08	0.20
Direction Network	0.00	0.00	0.06	0.00	0.00	0.06

TABLE XXX

Hypothesis testing results of the significance of frequencies of leadership-event sequences in Section 9.2.4.2. We reject H_0 at $\alpha = 0.01$. Each element in the table represents the percentage of the case when the test successfully rejects H_0.

Dyn. Type	**Type 1**			**Type 2**		
Model	**HM**	**DM**	**IC**	**HM**	**DM**	**IC**
Following Network	0.95	0.35	0.23	0.94	0.84	0.66
Direction Network	0.07	0.08	0.20	0.07	0.07	0.20

testing, the support distributions of the well-structure model, HM, are significantly different from the support distribution of the null model. The following networks also can be used to infer diagrams that are different from the rewiring diagrams than the direction networks based approach. However, in the simple type-2-dynamics datasets, our framework was able to infer diagrams that are more different from the null model compared to the complex type-1-dynamics case.

For the baboon dataset, we reported the information that we can retrieve from the dataset using our framework as a case study. Figure 43 shows the inferred diagram of leadership dynamics from our framework. Each row represents the node of leader sets of previous state and each column represents the next state. Each row label consists of baboon gender: 'M' or 'F', a set of frequent-leader IDs, and the support value of frequent-leader set. For example, in row 3 and column 2, the event that two female baboons F18 and F22 are leading their separate sub-groups concurrently can happen with the support 0.1 (out of all the coordination times). These two sub-groups have a chance to be merged together to a larger group lead by F18 with the probability 0.29. In 4th column ({F9}), we found that no matter what the previous sub-groups are, there is a high chance that the next group will be lead solely by the female baboon F9. In 4th row, F9 has the highest support (0.19), which means F9 (who happens to be the dominant female) often leads the troop, with the next highest support of 0.11 for the male baboon M3 (5th column, the alpha male). Lastly, at row 5 and column 4, if M3 and F9 are leading their separate sub-groups, then the two groups will be merged to a larger group lead by F9 with probability 0.63.

The hypothesis testing of the edge-weight distribution shows that the baboon's diagram is significantly different from the null model, with 100% of the time the tests successfully rejecting H_0. However, for the hypothesis testing of sequence-support distributions, the baboons' sequences of leadership dynamics are not significantly different from the rewired diagram. Only 5% of the times the tests successfully reject H_0. This indicates that while individual leaders identity is non-random and pairwise leadership transition patterns are significant, there are no

TABLE XXXI

Baboons' sequences of leadership dynamics that have the top-4 highest support

Baboon	**Sequences**	$\text{supp}_{path}(\mathcal{L}, P_{i,j})$
Seq. 1	{M11},{F9},{M3}	0.0354
Seq. 2	{M18},{F9},{M3}	0.0354
Seq. 3	{M18},{F9},{M22}	0.0354
Seq. 4	{M4},{F9},{M2}	0.0354

leadership sequences that often appear significantly within the baboon dataset. Nevertheless, Table XXXI shows baboons' sequences of leadership dynamics that have the top-4 highest support. This result is the evidence that F9 is an important individual who frequently leads the group.

These results show that our framework provides the opportunity for scientists to gain more insight into their datasets in order to generate scientific hypotheses, which might lead to important scientific discoveries (in this case, about the collective behavior and leadership dynamics of social animals).

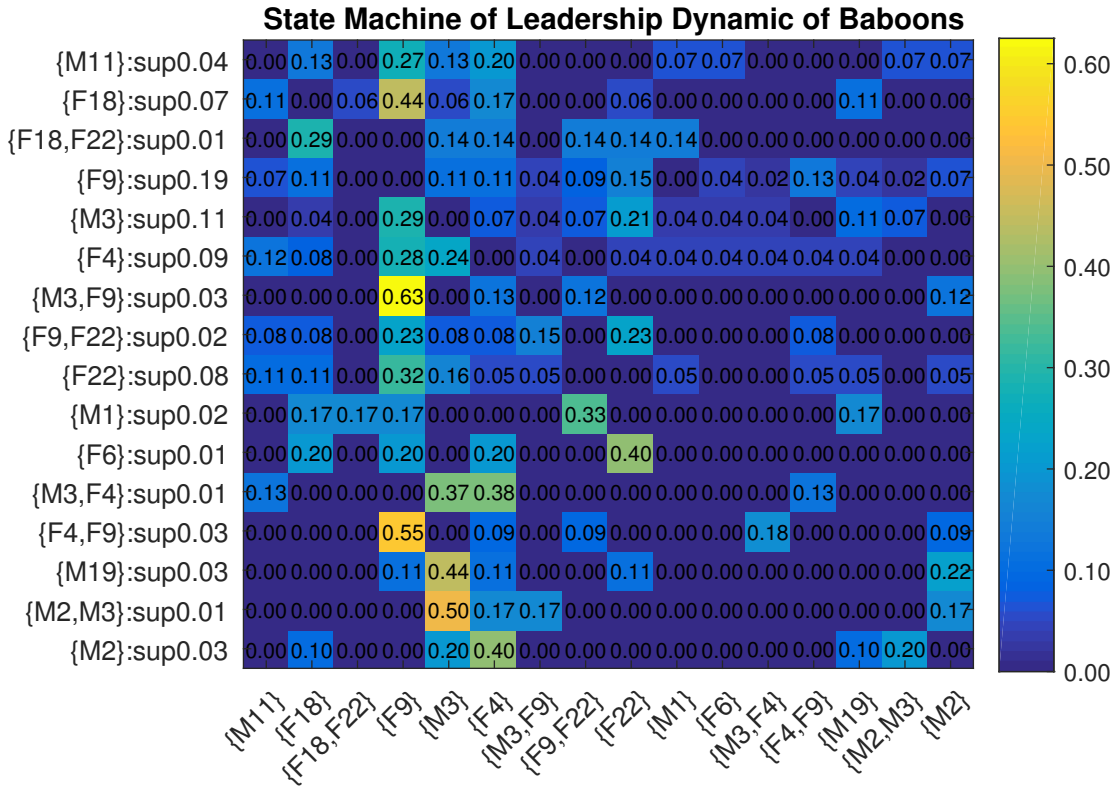


Figure 43. The inferred diagram of leadership dynamics of the baboon dataset from our framework. Each row represents the node of leader sets of previous state and each column represents the next state. Each row label consists of baboon gender: 'M' or 'F', a set of frequent-leader IDs, and the support value of frequent-leader set. For example, in row 3 and column 2, the event that two female baboons F18 and F22 are leading their separate sub-groups concurrently can happen with the support 0.1 (out of all the coordination times). These two sub-groups have a chance to be merged together to a larger group lead by F18 with the probability 0.29.

9.6 Conclusion

In chapter, we proposed a new approach to analyze time series of group movement data. We formalized a new computational problem, Mining Patterns of Leadership Dynamics, as well as proposed a framework as a solution of this problem. Our framework can be used to address several questions regarding leadership dynamics of group movement, such as 'what is the probability of having two sub-groups lead by i and j being merged together to be a larger group lead by k later?', 'what is the frequency of having i and k co-leading their sub-groups concurrently?', etc. We used the leadership inference framework, mFLICA [4], to infer the time series of leaders from movement datasets, then proposed an approach to mine and model frequent patterns of leadership dynamics using Hidden Markov Model. We evaluated our framework performance by using several simulated datasets that cover various types of leadership dynamics, as well as using the real-world dataset of baboon movement to demonstrate the application of our framework. There are no existing methods to address this problem, thus, we modified and extended the existing leadership inference framework to provide a non-trivial baseline. Our framework performed better than this baseline in all datasets. Moreover, we also proposed a method to perform statistical significance tests, comparing inferred diagram of leadership dynamics with our proposed null hypotheses to evaluate the significance of leadership-event order and frequencies of leadership-event sequences. Our framework opens the opportunities for scientists to generate scientific hypotheses that can be tested statistically regarding dynamics of leadership in movement data.

CHAPTER 10

CONCLUSION

10.1 Summary of contributions

This dissertation formalized concepts and problems of computational leadership inference in several aspects, as well as proposed computational frameworks to solve these problems.

In the first part, we explicitly formalized the definition of computational leaders as the "individuals who initiate collective patterns that everyone follows". We started with the formalization of a *following relation* among time series, which can be described intuitively as "two individuals performing the same sequence of actions (or generating time series values) with some fixed delay". We defined a *following network* with nodes representing time series and edges representing following relations between time series. Then, we formalized the concept of *coordination* as the time when "all individuals performing the same sequence of actions, at possibly varying delays". Afterwards, we defined an *initiator* or a leader as "an individual who first performs a sequence of actions, and all other individuals follow". After we had all the essential concepts, then we formalized Coordination Initiator Inference Problem, which is a problem of finding an initiator of coordination among time series. We proposed a framework, FLICA, to solve Coordination Initiator Inference Problem. FLICA performed well in both simulation and real-world datasets compared with the state-of-the-art approaches in the global leadership inference task, which has only one leader per dataset. Moreover, FLICA

also can handle several novel leadership-inference tasks that no existing methods can deal with, such as the local leadership inference (there is a different leader for a different coordination event), initiator of pre-coordination inference (there are different leaders at pre-coordination vs. coordination intervals), and leadership model selection.

In the second part, we extended the concept of coordinated activity from the first part to *multiple* coordinated activities which can occur simultaneously as well as proposed the framework, mFLICA, to analyze these complicated coordinated activities. We formalized the concept of a *faction* that is an induced subgraph of a following network such that a set of time series representing this subgraph is coordinated, and no another larger-coordinated subgraph that is a super set of this subgraph, which is a maximal coordinated subgraph. The concept of factions allows a set of time series to have multiple coordinated groups and faction initiators concurrently. We formalized Faction Initiator Inference Problem as a problem of finding factions and faction initiators from a given set of time series, as well as proposed mFLICA to solve this problem. While the previous work, FLICA, can detect only one coordination event and one initiator per time interval, mFLICA can detect more complicated leadership dynamics. We developed mFLICA based on the concept of factions that allows the framework to detect leader changing, group merging, and group splitting from time series. mFLICA performed well compared to our baseline in both simulated and fish datasets. We also demonstrated that mFLICA can detect the merging event of two subgroups in the baboon dataset.

In the third part, we proposed the concept of traits of leaders in time series based on the notion of convex hulls of movement features – velocity, position, and direction convex

hulls – to describe behavior traits of initiators. Specifically, if initiators always move first then others follow, then the initiators must step outside the group velocity convex hull from the previous time step frequently. If initiators always explore new area before others, then the initiators must step outside the previous group position convex hull frequently. Lastly, if everyone in the group always aligns its direction with initiators, then everyone else must step outside the previous group direction convex hull frequently but the initiators rarely step outside the direction convex hull. We provided a framework to infer traits of leaders of coordinated movement. The framework can infer ranking correlations of leadership ranking vs. convex hull ranking and can test the statistically significance of these behavior traits of initiators. We used baboon and fish datasets to demonstrate the ability of our framework. We found that baboon initiators do not move first, but typically explore new areas before others and the group aligns its direction with initiators. Fish initiators possess the same movement traits as baboon initiators' except that fish initiators also move before others.

In the fourth part, we proposed the new concept of inference of a coordination mechanism. We formalized a group coordinated activity as a process in which each individual uses some strategy to update its state to coordinate with the group. We defined a set of strategies that individuals may use to coordinated with the group as a set of *Coordination strategies*. We formalized Coordination Strategy Inference Problem as a problem of finding a set of coordination strategies that lead a group to coordinate its activity, given a set of possible strategies and a set of individual time series that has coordination. We proved that if every individual uses a strategy that guarantees to have a convergence, then the group will finally

reach coordination. We also proved that as long as all individuals in the group use any convergent strategy, then the group still reaches coordination even if some individuals use mixed convergent strategies. We proposed two convergent strategies: Hierarchical Model (HM) and Local Reversible Agreement System (LRA). For simplicity, we demonstrated our analysis in the movement context but our concept formalization can be applied in any context. In HM, an individual moves following a specific individual but there is one individual who moves toward a target state, whom everyone either directly or indirectly follows. In LRA, an individual moves following its physical neighbors but there is one individual who moves toward a target state. We proposed a framework to solve Coordination Strategy Inference Problem. In our setting, we have HM, LRA, and Auto Regressive (an individual moves toward the same direction as in its previous time steps) as our candidate strategies. We demonstrated our framework performance using simulated and animal datasets. Our framework performed well compared to our baseline. Moreover, our framework performed classification of group-level coordination models from time series better than FLICA framework, which is the-state-of-the-art approach for the task. Animal data experiments show that fish, unsurprisingly, follow their neighbors, while baboons have a preference to follow specific individuals (although that may depend on the time scale of analysis).

In the fifth part, we formalized the concept of a *frequent-leader set*. The frequent-leader set is a set of faction initiators that concurrently appear together frequently w.r.t. the support threshold. We defined a *transition probability of leader sets* as a probability that one frequent-leader set will transition to another frequent-leader set. Then we formalized Mining

Patterns of Leadership Dynamics as a problem of finding a set of frequent-leader sets and their transition probabilities, given a set of time series and a support threshold. We then modeled leadership dynamics of time series as a *diagram of leadership dynamics* that has a set of nodes representing a set of frequent-leader sets and a set of edges representing transition probabilities of these nodes. We proposed a framework to solve Mining Patterns of Leadership Dynamics. The framework adapts mFLICA and HMM as core components to infer a diagram of leadership dynamics from time series of movement data. The framework preformed well in simulated datasets compared to our non-trivial baseline. Moreover, we also proposed a method to perform statistical significance tests, comparing inferred diagram of leadership dynamics with our proposed null hypotheses to evaluate the significance of leadership-event order and frequencies of leadership-event sequences. We also demonstrated that our framework can report interesting patterns in the baboon dataset.

10.2 Assumptions and limitations of our formalization

First, a following relation that we defined here is only one of the possible definitions of a following relation in general. We defined a following relation as one of a pattern imitation, with a fixed time delay, between individuals. However, in real situations, a leader might influence followers to act differently or the followers might imitate their leader's actions with arbitrary and unfixed time delays.

Second, the concept of coordinated activity we defined also relies on the assumption that all possible pairs of individuals must follow the same collective pattern. However, our concept

of coordination does not cover the case when different individuals have been assigned to act differently to make the entire group reach collective goals.

Third, in this work, we mainly considered only time series of real values as inputs. That is, we never consider the time series of categorical values (e.g. walking, standing, running) in our analysis here.

Fourth, the concept of factions we defined has an assumption that all factions have disjoint members in the non-noise version of faction formalization. In contrast, different factions can have overlapping membership in the present of noise.

Lastly, in the work of inference of a coordination mechanism, we have a Markovian assumption that individuals update their states at time t based on their particular strategy that takes the parameters as the previous group states at time $t-1$. However, some individuals might update their states from the previous group states at some arbitrary time in the past.

10.3 Future directions

While we addressed several challenges in leadership inference from time series, there are still many directions that we can explore.

First, our computational leadership formalization as a process of pattern initiation is only one of the several possible definitions of leadership in general. As already mentioned, our leadership formalization can be used to describe only the case when followers imitate leaders' actions. To measure whether initiators truly cause the followers' actions, we should also explore other types of leadership definitions as well as concepts of causality inference from time series.

Second, we use only DTW as our main following relation measure that can find the lag patterns between time series. Nevertheless, different fields of study might have more specialized who-follows-whom measures. The question still remains regarding what measure is appropriate for the specific types of leadership systems.

Third, scalability is still an issue in our frameworks due to the time complexity of DTW. The alternative measures that can approximately measure following relations between time series should be explored.

Finally, all the work in this thesis is about how to infer leadership from time series but the question of why coordination exists in both natural and man-made systems still has an unclear answer. One of the possible ways to address this question is to use the game-theoretic approach to model leadership of coordination. If we can find that, under some assumptions, individuals gain more benefit to cooperate with others rather than solving tasks alone under leadership scenario, then we can use these assumptions to infer more properties of leadership from data and use them to answer scientific questions in both natural and man-made systems.

APPENDICES

Appendix A

COPYRIGHT PERMISSION STATEMENTS

COPYRIGHT TRANSFER AGREEMENT

In order for SIAM to include your paper in the 2018 SIAM International Conference on Data Mining proceedings, the following Copyright Transfer Agreement must be completed during the final paper upload process.

Title of Paper: Framework for Inferring Leadership Dynamics of Complex Movement from Time Series

Paper ID Number: 11

Copyright to this paper is hereby irrevocably assigned to SIAM for publication in the 2018 SIAM International Conference on Data Mining, May 3-5, 2018, at the San Diego Marriott Mission Valley, San Diego, California, USA. SIAM has sole use for distribution in all forms and media, such as microfilm and anthologies, except that the author(s) or, in the case of a "work made for hire," the employer will retain:

> The right to use all or part of the content of the paper in future works of the author(s), including author's teaching, technical collaborations, conference presentations, lectures, or other scholarly works and professional activities as well as to the extent the fair use provisions of the U.S. Copyright Act permit. If the copyright is granted to SIAM, then the proper notice of the SIAM's copyright should be provided.
>
> The right to post the final draft of the paper on noncommercial pre-print serves like arXiv.org.
>
> The right to post the final version of the paper on the author's personal web site and on the web server of the author's institution, provided the proper notice of the SIAM's copyright is included and that no separate or additional fees are collected for access to or distribution of the paper.
>
> The right to refuse permission to third parties to republish all or part of the paper or translation thereof.

It is affirmed neither this paper nor portions of it have been published elsewhere and a copyright transfer agreement has not been signed permitting the publication of a similar paper in a journal or elsewhere. For multi-author works, the signing author agrees to notify all co-authors of his/her action.

Transfer of Copyright to the Publisher
SIAM strongly recommends this option. This transfer of copyright provides SIAM the legal basis not only to publish and to distribute the work, but also to pursue infringements of copyright (such as plagiarism and other forms of unauthorized use) and to grant permissions for the legitimate uses of the work by third parties. This option should not be selected if the work was prepared by a government office or employee as part of his or her official duties.

[✓] The Author hereby irrevocably assigns, conveys and transfers the copyright to the Work to SIAM. SIAM shall have sole rights of distribution and/or publication of the work in all forms and media, throughout the world, except for those rights given to the Author above.

[__] The Author DOES NOT assign, convey and transfer the copyright to the Work to SIAM. Please list in whose name the copyright should appear here:
_____

Alternate Copyright
If you do not select either of the options above, an alternate copyright is required. Check below to submit an alternative copyright. Send proposed text for alternate copyright to erle@siam.org using the subject line "SIAM –SDM18 – Alternate Copyright – LAST NAME."

[__] I am submitting an alternative copyright.

Work Made for Hire
[__] Check here if signature is on behalf of employer in the event article is "work made for hire."

Previously Published
Have portions of this presentation been published elsewhere? If so, select "yes" and enclose appropriate credits and permissions to republish.
[__] Yes *(Enclose "Previously Published Credits and Permissions to Republish")*
[✓] No

ASSOCIATION FOR COMPUTING MACHINERY, INC. LICENSE TERMS AND CONDITIONS

Jun 27, 2018

This Agreement between Chainarong Amornbunchornvej ("You") and Association for Computing Machinery, Inc. ("Association for Computing Machinery, Inc.") consists of your license details and the terms and conditions provided by Association for Computing Machinery, Inc. and Copyright Clearance Center.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	4377400457474
License date	Jun 27, 2018
Licensed Content Publisher	Association for Computing Machinery, Inc.
Licensed Content Publication	ACM Transactions on Knowledge Discovery from Data
Licensed Content Title	Coordination Event Detection and Initiator Identification in Time Series Data
Licensed Content Author	Chainarong Amornbunchornvej, et al
Licensed Content Date	Jun 27, 2018
Licensed Content Volume	12
Licensed Content Issue	5
Volume number	12
Issue number	5
Type of Use	Thesis/Dissertation
Requestor type	Author of this ACM article
Is reuse in the author's own new work?	Yes
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Inference of Leadership of Coordinated Activity in Time Series
Expected completion date	Dec 2018
Estimated size (pages)	250
Requestor Location	Chainarong Amornbunchornvej 1040 W Granville AVE APT 719 CHICAGO, IL 60660 United States Attn: Chainarong Amornbunchornvej
Billing Type	Credit Card
Credit card info	Visa ending in 8473
Credit card expiration	04/2019
Total	8.00 USD
Terms and Conditions	

ASSOCIATION FOR COMPUTING MACHINERY, INC. LICENSE TERMS AND CONDITIONS

Jun 14, 2018

This Agreement between Chainarong Amornbunchornvej ("You") and Association for Computing Machinery, Inc. ("Association for Computing Machinery, Inc.") consists of your license details and the terms and conditions provided by Association for Computing Machinery, Inc. and Copyright Clearance Center.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	4366550974192
License date	Jun 12, 2018
Licensed Content Publisher	Association for Computing Machinery, Inc.
Licensed Content Publication	Proceedings
Licensed Content Title	Identifying Traits of Leaders in Movement Initiation
Licensed Content Author	Chainarong Amornbunchornvej, et al
Licensed Content Date	Jul 31, 2017
Type of Use	Thesis/Dissertation
Requestor type	Author of this ACM article
Is reuse in the author's own new work?	Yes
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Inference of Leadership of Coordinated Activity in Time Series
Expected completion date	Dec 2018
Estimated size (pages)	250
Requestor Location	Chainarong Amornbunchornvej 1040 W Granville AVE APT 719 CHICAGO, IL 60660 United States Attn: Chainarong Amornbunchornvej
Billing Type	Credit Card
Credit card info	Visa ending in 8473
Credit card expiration	04/2019
Total	8.00 USD
Terms and Conditions	

CITED LITERATURE

1. Amornbunchornvej, C., Brugere, I., Strandburg-Peshkin, A., Farine, D., Crofoot, M. C., and Berger-Wolf, T. Y.: Flica: A framework for leader identification in coordinated activity. *arXiv preprint arXiv:1603.01570* , 2016.

2. Amornbunchornvej, C., Brugere, I., Strandburg-Peshkin, A., Farine, D., Crofoot, M. C., and Berger-Wolf, T. Y.: Coordination event detection and initiator identification in time series data. In *SIGKDD Workshop on Mining and Learning from Time Series (MiLeTS 2017)* , pages 1–9, August 2017.

3. Amornbunchornvej, C., Brugere, I., Strandburg-Peshkin, A., Farine, D., Crofoot, M. C., and Berger-Wolf, T. Y.: Coordination event detection and initiator identification in time series data. *ACM Trans. Knowl. Discov. Data* , 12(5):1–33, 7 2018.

4. Amornbunchornvej, C. and Berger-Wolf, T.: Framework for inferring leadership dynamics of complex movement from time series. In *Proceedings of the 2018 SIAM International Conference on Data Mining* , pages 549–557. SIAM, 2018.

5. Amornbunchornvej, C., Crofoot, M. C., and Berger-Wolf, T. Y.: Identifying traits of leaders in movement initiation. In *ASONAM'17* , pages 660–666, July 2017.

6. Crofoot, M. C., Kays, R. W., and Wikelski, M.: Data from: Shared decision-making drives collective movement in wild baboons, 2015.

7. Strandburg-Peshkin, A., Farine, D. R., Couzin, I. D., and Crofoot, M. C.: Shared decision-making drives collective movement in wild baboons. *Science* , 348(6241):1358–1361, 2015.

8. Dyer, J. R., Johansson, A., Helbing, D., Couzin, I. D., and Krause, J.: Leadership, consensus decision making and collective behaviour in humans. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* , 364(1518):781–789, 2009.

9. Stewart, J. C. and Scott, J.: Lack of correlation between leadership and dominance relationships in a herd of goats. *Journal of comparative and physiological psychology* , 40(4):255, 1947.

10. Hogg, M. A.: A social identity theory of leadership. *Personality and social psychology review* , 5(3):184–200, 2001.

11. Glowacki, L. and von Rueden, C.: Leadership solves collective action problems in small-scale societies. *Phil. Trans. R. Soc. B* , 370(1683):20150010, 2015.

12. Weinstein, P. and Maelzer, D. A.: Leadership behaviour in sawfly larvae perga dorsalis (hymenoptera: Pergidae). *Oikos* , 79(3):450–455, 1997.

13. Heinsohn, R. and Packer, C.: Complex cooperative strategies in group-territorial african lions. *Science* , 269(5228):1260–1262, 1995.

14. Boydston, E. E., Morelli, T. L., and Holekamp, K. E.: Sex differences in territorial behavior exhibited by the spotted hyena (hyaenidae, crocuta crocuta). *Ethology* , 107(5):369–385, 2001.

15. Mares, R., Young, A. J., and Clutton-Brock, T. H.: Individual contributions to territory defence in a cooperative breeder: weighing up the benefits and costs. *Proceedings of the Royal Society of London B: Biological Sciences* , 279(1744):3989–3995, 2012.

16. Gilby, I. C., Machanda, Z. P., Mjungu, D. C., Rosen, J., Muller, M. N., Pusey, A. E., and Wrangham, R. W.: 'impact hunters' catalyse cooperative hunting in two wild chimpanzee communities. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* , 370(1683), 2015.

17. Glowacki, L. and von Rueden, C.: Leadership solves collective action problems in small-scale societies. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* , 370(1683), 2015.

18. Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. A.: Effective leadership and decision-making in animal groups on the move. *Nature* , 433(7025):513–516, 2005.

19. Hooper, P. L., Kaplan, H. S., and Boone, J. L.: A theory of leadership in human cooperative groups. *Journal of Theoretical Biology* , 265(4):633 – 646, 2010.

20. Malone, T. W. and Crowston, K.: The interdisciplinary study of coordination. *ACM Computing Surveys (CSUR)* , 26(1):87–119, 1994.

21. Krause, J., Hoare, D., Krause, S., Hemelrijk, C., and Rubenstein, D.: Leadership in fish shoals. *Fish and Fisheries* , 1(1):82–89, 2000.

22. Wilson, E. O.: Sociobiology: the new synthesis. *Philosophy of Biology: An Anthology* , page 339, 2009.

23. Stueckle, S. and Zinner, D.: To follow or not to follow: decision making and leadership during the morning departure in chacma baboons. *Animal Behaviour* , 75(6):1995 – 2004, 2008.

24. Petit, O. and Bon, R.: Decision-making processes: The case of collective movements. *Behavioural Processes* , 84(3):635 – 647, 2010. Special section: Collective movements.

25. Wu, S. and Sun, Q.: Computer simulation of leadership, consensus decision making and collective behaviour in humans. *PLOS ONE* , 9(1):1–12, 01 2014.

26. Brent, L. J., Franks, D. W., Foster, E. A., Balcomb, K. C., Cant, M. A., and Croft, D. P.: Ecological knowledge, leadership, and the evolution of menopause in killer whales. *Current Biology* , 25(6):746 – 750, 2015.

27. Solera, F., Calderara, S., and Cucchiara, R.: Learning to identify leaders in crowd. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* , pages 43–48, 2015.

28. Pham, H. and Shahabi, C.: Spatial influence - measuring followship in the real world. In *ICDE16* , pages 529–540, May 2016.

29. Lusseau, D. and Conradt, L.: The emergence of unshared consensus decisions in bottlenose dolphins. *Behavioral Ecology and Sociobiology* , 63(7):1067–1077, 2009.

30. Andersson, M., Gudmundsson, J., Laube, P., and Wolle, T.: Reporting leaders and followers among trajectories of moving point objects. *GeoInformatica* , 12(4):497–528, 2008.

31. Goyal, A., Bonchi, F., and Lakshmanan, L. V.: Discovering leaders from community actions. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management* , CIKM '08, pages 499–508, New York, NY, USA, 2008. ACM.

32. Bakshy, E., Hofman, J. M., Mason, W. A., and Watts, D. J.: Everyone's an influencer: Quantifying influence on twitter. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining* , WSDM '11, pages 65–74, New York, NY, USA, 2011. ACM.

33. Kjargaard, M. B., Blunck, H., Wustenberg, M., Gronbask, K., Wirz, M., Roggen, D., and Troster, G.: Time-lag method for detecting following and leadership behavior of pedestrians from mobile sensing data. In *Proceedings of the IEEE PerCom* , pages 56–64. IEEE, 2013.

34. Sun, J. and Tang, J.: *A Survey of Models and Algorithms for Social Influence Analysis* , pages 177–214. Boston, MA, Springer US, 2011.

35. Benkert, M., Gudmundsson, J., Hübner, F., and Wolle, T.: *Reporting flock patterns* . Springer, 2006.

36. Kempe, D., Kleinberg, J., and Tardos, É.: Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD* , pages 137–146. ACM, 2003.

37. Goyal, A., Bonchi, F., and Lakshmanan, L. V.: Learning influence probabilities in social networks. In *Proceedings of the third ACM international conference on Web search and data mining* , pages 241–250. ACM, 2010.

38. Sánchez-Amaro, A., Duguid, S., Call, J., and Tomasello, M.: Chimpanzees, bonobos and children successfully coordinate in conflict situations. *Proceedings of the Royal Society of London B: Biological Sciences* , 284(1856), 2017.

39. Beek, P., Peper, C., and Stegeman, D.: Dynamical models of movement coordination. *Human Movement Science* , 14(4):573 – 608, 1995.

40. Cao, Y., Yu, W., Ren, W., and Chen, G.: An overview of recent progress in the study of distributed multi-agent coordination. *IEEE Transactions on Industrial Informatics* , 9(1):427–438, Feb 2013.

41. Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. A.: Effective leadership and decision-making in animal groups on the move. *Nature* , 433(7025):513–516, 2005.

42. Smith, J. E., Estrada, J. R., Richards, H. R., Dawes, S. E., Mitsos, K., and Holekamp, K. E.: Collective movements, leadership and consensus costs at reunions in spotted hyaenas. *Animal Behaviour* , 105:187–200, 2015.

43. Stueckle, S. and Zinner, D.: To follow or not to follow: decision making and leadership during the morning departure in chacma baboons. *Animal Behaviour* , 75(6):1995–2004, 2008.

44. Yu, C.-H., Werfel, J., and Nagpal, R.: Collective decision-making in multi-agent systems by implicit leadership. In *AAMAS'10* , pages 1189–1196, May 2010.

45. Lewis, F. L., Zhang, H., Hengster-Movric, K., and Das, A.: *Cooperative control of multi-agent systems: optimal and adaptive design approaches* . Springer Science & Business Media, 2013.

46. Strandburg-Peshkin, A. and et al.: Visual sensory networks and effective information transfer in animal groups. *Current Biology* , 23(17):R709–R711, 2013.

47. Chazelle, B.: The total s-energy of a multiagent system. *SIAM Journal on Control and Optimization* , 49(4):1680–1706, 2011.

48. He, X. and Kempe, D.: Robust influence maximization. In *Proceedings of the ninth ACM SIGKDD* , pages 1–10. ACM, 2016.

49. Farine, D. R., Strandburg-Peshkin, A., Couzin, I. D., Berger-Wolf, T. Y., and Crofoot, M. C.: Individual variation in local interaction rules can explain emergent patterns of spatial organization in wild baboons. *Proceedings of the Royal Society of London B: Biological Sciences* , 284(1853), 2017.

50. Farine, D. R., Strandburg-Peshkin, A., Berger-Wolf, T., Ziebart, B., Brugere, I., Li, J., and Crofoot, M. C.: Both nearest neighbours and long-term affiliates predict individual locations during collective movement in wild baboons. *Scientific reports* , 6:27704, 2016.

51. Gautrais, J., Ginelli, F., Fournier, R., Blanco, S., Soria, M., Chat, H., and Theraulaz, G.: Deciphering interactions in moving animal groups. *PLOS Computational Biology* , 8(9):1–11, 09 2012.

52. Herbert-Read, J. E., Perna, A., Mann, R. P., Schaerf, T. M., Sumpter, D. J. T., and Ward, A. J. W.: Inferring the rules of interaction of shoaling fish. *Proceedings of the National Academy of Sciences* , 108(46):18726–18731, 2011.

53. Katz, Y., Tunstrøm, K., Ioannou, C. C., Huepe, C., and Couzin, I. D.: Inferring the structure and dynamics of interactions in schooling fish. *Proceedings of the National Academy of Sciences* , 108(46):18720–18725, 2011.

54. Mann, R. P., Perna, A., Strmbom, D., Garnett, R., Herbert-Read, J. E., Sumpter, D. J. T., and Ward, A. J. W.: Multi-scale inference of interaction rules in animal groups using bayesian model selection. *PLOS Computational Biology* , 9(3):1–13, 03 2013.

55. Langrock, R., Hopcraft, J. G. C., Blackwell, P. G., Goodall, V., King, R., Niu, M., Patterson, T. A., Pedersen, M. W., Skarin, A., and Schick, R. S.: Modelling group dynamic animal movement. *Methods in Ecology and Evolution* , 5(2):190–199, 2014.

56. Li, Z., Lee, J.-G., Li, X., and Han, J.: Incremental clustering for trajectories. In *Database Systems for Advanced Applications* , eds. H. Kitagawa, Y. Ishikawa, Q. Li, and C. Watanabe, pages 32–46, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

57. Spiliopoulou, M., Ntoutsi, I., Theodoridis, Y., and Schult, R.: Monic: Modeling and monitoring cluster transitions. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* , KDD '06, pages 706–711, New York, NY, USA, 2006. ACM.

58. Amornbunchornvej, C. and Berger-Wolf, T. Y.: Mining and modeling complex leadership dynamics of movement data. In *ASONAM'18* , pages x–x, August 2018.

59. Brown, C. and Irving, E.: Individual personality traits influence group exploration in a feral guppy population. *Behavioral Ecology* , 25(1):95, 2014.

60. Laube, P., van Kreveld, M., and Imfeld, S.: Finding remo — detecting relative motion patterns in geospatial lifelines. In *Developments in Spatial Data Handling* , pages 201–215, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

61. Agrawal, R., Imieliński, T., and Swami, A.: Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data* , SIGMOD '93, pages 207–216, New York, NY, USA, 1993. ACM.

62. Han, J., Cheng, H., Xin, D., and Yan, X.: Frequent pattern mining: current status and future directions. *Data Mining and Knowledge Discovery* , 15(1):55–86, Aug 2007.

63. Aggarwal, C. C. and Han, J.: *Frequent pattern mining* . Springer, 2014.

64. Rabiner, L. R.: A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* , 77(2):257–286, Feb 1989.

65. Ben Dushnik, E. W. M.: Partially ordered sets. *American Journal of Mathematics* , 63(3):600–610, 1941.

66. Page, L., Brin, S., Motwani, R., and Winograd, T.: The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999.

67. Li, J., Asif, K., Wang, H., Ziebart, B. D., and Berger-Wolf, T. Y.: Adversarial sequence tagging. In *IJCAI* , pages 1690–1696, 2016.

68. Sakoe, H. and Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing* , 26(1):43–49, Feb 1978.

69. Shokoohi-Yekta, M., Wang, J., and Keogh, E.: On the non-trivial generalization of dynamic time warping to the multi-dimensional case. In *SDM'15* , pages 289–297, 2015.

70. Holme, P.: *Temporal networks* . Springer, 2014.

71. Kendall, M. G.: A new measure of rank correlation. *Biometrika* , 30(1/2):81–93, 1938.

72. Liu, Y., Bahadori, T., and Li, H.: Sparse-gev: Sparse latent space model for multivariate extreme value time serie modeling. *ICML* , 2012.

73. Sulo, R., Berger-Wolf, T., and Grossman, R.: Meaningful selection of temporal resolution for dynamic networks. In *Proceedings of the Eighth Workshop on Mining and Learning with Graphs* , pages 127–136. ACM, 2010.

74. Ho, T. K.: The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence* , 20(8):832–844, 1998.

75. Amornbunchornvej, C.: Flica: A framework for leader identification in coordinated activity codes. `https://github.com/CompBioUIC/FLICA`, 2017. Accessed: 2017-08-30.

76. Sakoe, H. and Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing* , 26(1):43–49, 1978.

77. Will, T. E.: Flock leadership: Understanding and influencing emergent collective behavior. *The Leadership Quarterly* , 27(2):261–279, 2016.

78. mFLICA: code and supplementary. `https://github.com/CompBioUIC/MFLICA`. Accessed: 2017-12-19.

79. Byrd, R. H., Hribar, M. E., and Nocedal, J.: An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization* , 9(4):877–900, 1999.

80. Coordination model selection framework code and data. `https://github.com/CompBioUIC/CoordinationModelSelectionFramework`. Accessed: 2018-02-06.

81. Jelinek, F., Bahl, L., and Mercer, R.: Design of a linguistic statistical decoder for the recognition of continuous speech. *IEEE Transactions on Information Theory* , 21(3):250–256, 1975.

82. Massey Jr., F. J.: The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association* , 46(253):68–78, 1951.

83. Wilcoxon, F.: Individual comparisons by ranking methods. *Biometrics Bulletin* , 1(6):80–83, 1945.

84. Kruskal, W. H. and Wallis, W. A.: Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* , 47(260):583–621, 1952.

VITA

NAME: Chainarong Amornbunchornvej

EDUCATION: Ph.D., Computer Science, University of Illinois at Chicago, Chicago, Illinois, 2018.

M.Eng., Telecommunications Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, 2013

B.Eng. (honor), Computer Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, 2011

ACADEMIC EXPERIENCE: Research Assistant, Computational Population Biology Lab, Department of Computer Science, University of Illinois at Chicago, 2015 - 2018.

Teaching Assistant, Department of Computer Science, University of Illinois at Chicago:

- Computer Algorithm I, Spring 2016 and Fall 2017.
- Secure Computer Systems, Fall 2015

PROFESSIONAL EXPERIENCE: Software Developer, Draycir Co, Ltd., Bangkok, Thailand, 2014

Research Intern, National Center for Genetic Engineering and Biotechnology, Pathum Thani, Thailand , 2011 - 2013

PROFESSIONAL MEMBERSHIP: Society for Industrial and Applied Mathematics (SIAM)

Association for Computing Machinery (ACM)

HONORS: Royal Thai Government Scholarship for Studying Doctoral Program, 2014 - 2018

Junior Science Talent Project long-term Scholarship by National Science and Technology Development Agency, Thailand, 2007 - 2012

PUBLICATIONS

1. Amornbunchornvej, C., and Tanya Y. Berger-Wolf. "Mining and Modeling Complex Leadership Dynamics of Movement data." In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM18), ACM, 2018.

2. Amornbunchornvej, C., Ivan Brugere, Ariana Strandburg-Peshkin, Damien Farine, Margaret C. Crofoot, and Tanya Y. Berger-Wolf. "Coordination Event Detection and Initiator Identification in Time Series Data." ACM Transactions on Knowledge Discovery from Data (TKDD) 12, no. 5 (June 2018): 53.

3. Amornbunchornvej, C., and Tanya Y. Berger-Wolf. "Framework for Inferring Leadership Dynamics of Complex Movement from Time Series." In Proceedings of the 2018 SIAM International Conference on Data Mining (SDM18), Society for Industrial and Applied Mathematics, 2018.

4. Amornbunchornvej, C., Margaret C. Crofoot, and Tanya Y. Berger-Wolf. "Identifying Traits of Leaders in Movement Initiation." In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM17), pp. 660-667. ACM, 2017.

5. Amornbunchornvej, C., Ivan Brugere, Ariana Strandburg-Peshkin, Damien R. Farine, Margaret C. Crofoot, and Tanya Y. Berger-Wolf. "Coordination Event Detection and Initiator Identification in Time Series Data." 3RD SIGKDD workshop on mining and learning from time series, 2017.

6. Amornbunchornvej, C., Ivan Brugere, Ariana Strandburg-Peshkin, Damien Farine, Margaret C. Crofoot, and Tanya Y. Berger-Wolf. "FLICA: a framework for leader identification in coordinated activity." arXiv preprint arXiv:1603.01570 (2016).

7. Limpiti, T., Chainarong Amornbunchornvej, Apichart Intarapanich, Anunchai Assawamakin, and Sissades Tongsima. "iNJclust: iterative neighbor-joining tree clustering framework for inferring population structure." IEEE/ACM transactions on computational biology and bioinformatics 11, no. 5 (2014): 903-914.

8. Amornbunchornvej, C. "Iterative Clustering using Hierarchical Tree for High-dimensional Genotypic Data." Masters thesis, King Mongkut's Institute of Technology Ladkrabang, 2013.

9. Amornbunchornvej, C., Tulaya Limpiti, Anunchai Assawamakin, Apichart Intarapanich, and Sissades Tongsima. “Iterative neighbor-joining tree clustering algorithm for genotypic data.” In Pattern Recognition (ICPR), 2012 21st International Conference on, pp. 1827-1830. IEEE, 2012.

10. Amornbunchornvej, C., T. Limpiti, A. Assawamakin, A. Intarapanich, and S. Tongsima. “Improved iterative pruning principal component analysis with graph-theoretic hierarchical clustering.” In Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2012 9th International Conference on, pp. 1-4. IEEE, 2012.