Computing Consensus Curves

Sergey Pupyrev
University of Arizona

Joint work with Livio De La Cruz, Stephen Kobourov, Paul Shen, and Sankar Veeramoni
Analyzing Insect Colonies

Motivation - discovering behavioral patterns in ant colonies:

- how often ants communicate
- what roles do ants play in a colony
- how does interaction and communication affect the success or failure of a colony
Analyzing Insect Colonies

Motivation - discovering behavioral patterns in ant colonies:

- how often ants communicate
- what roles do ants play in a colony
- how does interaction and communication affect the success or failure of a colony

tracking the ants’ motion is needed!
Analyzing Insect Colonies
Current Approaches

- gluing bar codes on to ants
Current Approaches

- gluing bar codes on to ants
- computer vision techniques

(initialization)  (frame 460)
Current Approaches

- gluing bar codes on to ants
- computer vision techniques
  - requires high-resolution video
  - accuracy $\approx 80\%$
  - poor results on long videos

initialization  
frame 460
Citizen Science

- online game at http://angryants.cs.arizona.edu
Citizen Science

- online game at http://angryants.cs.arizona.edu
- clicks on top of ants induce trajectories
Citizen Science

- online game at http://angryants.cs.arizona.edu
- clicks on top of ants induce trajectories

The Goal

Computing consensus trajectories from many (possibly inaccurate) trajectories submitted by citizen scientists.
Local Approach
Local Approach

set of $m$ trajectories for $k$ ants ($m > k$)

$$
\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)
$$

$$
\ldots
$$

$$
\tau_m = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)
$$
Local Approach

set of $m$ trajectories for $k$ ants ($m > k$)

$\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)$

$\tau_m = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)$

consider trajectories for $i$-th ant

1. Local Mean: $\tau = \{(x_{avg}, y_{avg}, t_i)\}$
Local Approach

set of \( m \) trajectories for \( k \) ants \((m > k)\)

\[
\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)
\]

\[
\cdots
\]

\[
\tau_m = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)
\]

consider trajectories for \( i \)-th ant

1. Local Mean:  
\[
\tau = \{(x_{avg}, y_{avg}, t_i)\}
\]
Local Approach

set of $m$ trajectories for $k$ ants ($m > k$)

$$
\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)
$$

$$
\ldots
$$

$$
\tau_m = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)
$$

consider trajectories for $i$-th ant

1. Local Mean: $\tau = \{ (x_{avg}, y_{avg}, t_i) \}$
Local Approach

set of $m$ trajectories for $k$ ants ($m > k$)

$\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)$

$\tau_m = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)$

consider trajectories for $i$-th ant

1. Local Mean: $\tau = \{(x_{avg}, y_{avg}, t_i)\}$

2. Local Median: $\tau = \{(x_{med}, y_{med}, t_i)\}$
Local Approach

set of \( m \) trajectories for \( k \) ants \((m > k)\)

\[
\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)
\]

\[
\ldots
\]

\[
\tau_m = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)
\]

consider trajectories for \( i \)-th ant

1. Local Mean: \( \tau = \{(x_{\text{avg}}, y_{\text{avg}}, t_i)\} \)

2. Local Median: \( \tau = \{(x_{\text{med}}, y_{\text{med}}, t_i)\} \)
Local Approach

set of $m$ trajectories for $k$ ants ($m > k$)

$$\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)$$

$$\tau_2 = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)$$

consider trajectories for $i$-th ant

1. Local Mean: $\tau = \{(x_{avg}, y_{avg}, t_i)\}$

2. Local Median: $\tau = \{(x_{med}, y_{med}, t_i)\}$
Local Approach

set of $m$ trajectories for $k$ ants ($m > k$)

$$
\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)
$$

$$
\ldots
$$

$$
\tau_m = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)
$$

consider trajectories for $i$-th ant

1. Local Mean: $\tau = \{(x_{avg}, y_{avg}, t_i)\}$
2. Local Median: $\tau = \{(x_{med}, y_{med}, t_i)\}$
3. Local Fréchet:
4. Homotopy Median: Wiratma et al., GIS’11
5. Majority Median: Buchin et al., Algorithmica’13
Local Approach

set of $m$ trajectories for $k$ ants ($m > k$)

$$\tau_1 = (x_{11}, y_{11}, t_1), \ldots, (x_{1T}, y_{1T}, t_T)$$

$$\ldots$$

$$\tau_m = (x_{m1}, y_{m1}, t_1), \ldots, (x_{mT}, y_{mT}, t_T)$$

consider trajectories for $i$-th ant

1. Local Mean: $\tau = \{(x_{avg}, y_{avg}, t_i)\}$

2. Local Median: $\tau = \{(x_{med}, y_{med}, t_i)\}$

3. Local Fréchet:

4. Homotopy Median: Wiratma et al., GIS’11

5. Majority Median: Buchin et al., Algorithmica’13
Local Approach
Local Approach

two ants on top of each other
Local Approach

two ants on top of each other
valid pieces of trajectories
Global Approach
Global Approach
Global Approach

clustering points
Global Approach

- Clustering points
- Graph construction
Global Approach

clustering points

3 weights on edges: “red”, “blue”, “green”

graph construction
Global Approach

clustering points

3 weights on edges: “red”, “blue”, “green”

graph construction
Simultaneous Consensus Paths

Input

- DAG with $k$ sources and $k$ sinks
Simultaneous Consensus Paths

Input

- DAG with $k$ sources and $k$ sinks
- $\text{deg}^{in} = \text{deg}^{out}$ for internal nodes
Simultaneous Consensus Paths

Input

- DAG with $k$ sources and $k$ sinks
- $\text{deg}^{in} = \text{deg}^{out}$ for internal nodes
- edge-weights $w^i(e)$ for every color $1 \leq i \leq k$
Simultaneous Consensus Paths

Input

- DAG with $k$ sources and $k$ sinks
- $\deg^{in} = \deg^{out}$ for internal nodes
- edge-weights $w^i(e)$ for every color $1 \leq i \leq k$

Output

- $k$ edge-disjoint paths
Simultaneous Consensus Paths

Input

- DAG with \( k \) sources and \( k \) sinks
- \( \text{deg}^{in} = \text{deg}^{out} \) for internal nodes
- edge-weights \( w^i(e) \) for every color \( 1 \leq i \leq k \)

Output

- \( k \) edge-disjoint paths
- maximize
  \[
  \sum_{i=1}^{k} \sum_{e \in P_i} w^i(e)
  \]
Simultaneous Consensus Paths

Input

- DAG with $k$ sources and $k$ sinks
- $\text{deg}^{in} = \text{deg}^{out}$ for internal nodes
- edge-weights $w^i(e)$ for every color $1 \leq i \leq k$

Output

- $k$ edge-disjoint paths
- maximize $\sum_{i=1}^{k} \sum_{e \in P_i} w^i(e)$

Theorem 1

The SCP problem is NP-hard, even when restricted to planar grid graphs
Simultaneous Consensus Paths

**Theorem 1**
The SCP problem is NP-hard, even when restricted to planar grid graphs

**Theorem 2**
The SCP problem can be solved optimally in $O(|E| + k!|V|)$ time

- $\text{deg}^{in} = \text{deg}^{out}$ for internal nodes
- edge-weights $w^i(e)$ for every color $1 \leq i \leq k$
- maximize $\sum_{i=1}^{k} \sum_{e \in P_i} w^i(e)$
Simultaneous Consensus Paths

Practical heuristics:
Simultaneous Consensus Paths

Practical heuristics:
- the greedy algorithm
Simultaneous Consensus Paths

Practical heuristics:
- the greedy algorithm

1. find the heaviest path
Simultaneous Consensus Paths

Practical heuristics:

- the greedy algorithm

1. find the heaviest path
2. remove it

![Graph with weights]

- \( w^G_{10} \)
- \( w^B_{10} \)
- \( w^G_4 \)
- \( w^B_3 \)
- \( w^G_2 \)
- \( w^B_1 \)
- \( w^G_2 \)
- \( w^B_3 \)
- \( w^G_6 \)
- \( w^B_7 \)
- \( w^G_6 \)
- \( w^B_7 \)
- \( w^G_6 \)
- \( w^B_5 \)
Simultaneous Consensus Paths

Practical heuristics:

- the greedy algorithm

1. find the heaviest path
2. remove it
Simultaneous Consensus Paths

Practical heuristics:
- the greedy algorithm

1. find the heaviest path
2. remove it
Simultaneous Consensus Paths

Practical heuristics:

- the greedy algorithm

1. find the heaviest path
2. remove it
Simultaneous Consensus Paths

Practical heuristics:

- the greedy algorithm

1. find the heaviest path
2. remove it
Simultaneous Consensus Paths

Practical heuristics:

- the greedy algorithm

1. find the heaviest path
2. remove it
Simultaneous Consensus Paths

Practical heuristics:
- the greedy algorithm
- integer linear programming (ILP)

maximize \[ \sum_i \sum_e w^i(e)x_e^i \]
subject to \[ \sum_i x_e^i = 1 \quad \forall e \in E \]
\[ \sum_{uv} x_{uv}^i = \sum_{vw} x_{vw}^i \quad \forall v \in V \setminus \{s_1, t_1, \ldots, s_k, t_k\} \]
\[ \sum_v x_{sv}^i = 1 \quad \forall 1 \leq i \leq k \]
\[ x_e^i \in \{0, 1\} \quad \forall e \in E, 1 \leq i \leq k \]
Simultaneous Consensus Paths

Practical heuristics:
- the greedy algorithm
- integer linear programming (ILP)
- randomized rounding

Maximize

\[ \sum_i \sum_e w^i(e) x^i_e \]

Subject to

\[ \sum_i x^i_e = 1 \quad \forall e \in E \]
\[ \sum_{uv} x^i_{uv} = \sum_{vw} x^i_{vw} \quad \forall v \in V \setminus \{s_1, t_1, \ldots, s_k, t_k\} \]
\[ \sum_v x^i_{s_i,v} = 1 \quad \forall 1 \leq i \leq k \]
\[ x^i_e \in \{0, 1\} \quad \forall e \in E, 1 \leq i \leq k \]

Probability of choosing \( i \)-th ant

Relaxed to \( 0 \leq x^i_e \leq 1 \)
Experiments

Dataset:

- Temnothorax rugatulus colony, 50 ants
- $\approx 5$ minutes video, 100 timeframes, 252 contributed trajectories
- analyzed by Shin et al., WACV’11-12
Experiments

Dataset:
- **Temnothorax rugatulus colony, 50 ants**
- ≈ 5 minutes video, 100 timeframes, 252 contributed trajectories
- analyzed by Shin et al., WACV’11-12

Automated image-processing tracking system:
- *manually* painted ants
- took ≈ 2.5 hours to process
Experiments

Dataset:
- Temnothorax rugatulus colony, 50 ants
- ≈ 5 minutes video, 100 timeframes, 252 contributed trajectories
- analyzed by Shin et al., WACV’11-12

Automated image-processing tracking system:
- manually painted ants
- took ≈ 2.5 hours to process

Ground truth:
- created by manually inspecting the automated solution and fixing errors
## Experiments

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Worst RMSE</th>
<th>Average RMSE</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated Solution</td>
<td>95.308</td>
<td>9.593</td>
<td>160 min</td>
</tr>
<tr>
<td>Local Mean</td>
<td>105.271</td>
<td>12.531</td>
<td>&lt; 100 ms</td>
</tr>
<tr>
<td>Local Median</td>
<td>112.741</td>
<td>9.801</td>
<td>&lt; 100 ms</td>
</tr>
<tr>
<td>Local Fréchet</td>
<td>127.104</td>
<td>15.562</td>
<td>1.2 sec</td>
</tr>
<tr>
<td>Homotopy Median</td>
<td>146.267</td>
<td>20.244</td>
<td>8.2 sec</td>
</tr>
<tr>
<td>Buffer Median</td>
<td>171.556</td>
<td>23.998</td>
<td>9.7 sec</td>
</tr>
<tr>
<td>Global ILP</td>
<td>20.588</td>
<td>8.716</td>
<td>34 sec</td>
</tr>
<tr>
<td>Global Greedy</td>
<td>24.820</td>
<td>8.900</td>
<td>0.2 sec</td>
</tr>
</tbody>
</table>

root-mean-square error (RMSE): \[
\sqrt{\frac{1}{n} \sum_{i=1}^{T} \| \tau(t_i) - \tau^{OPT}(t_i) \|^2}
\]

\(\approx 60 \text{ pixels}\)
Experiments

Synthetic dataset:

- “realistic” trajectories (follows models of ant behaviour by Scheidler et al., Depickére et al.)

- varying #ants, #timestamps, #trajectories per ant, probability of making a mistake
Experiments

Synthetic dataset:

- “realistic” trajectories (follows models of ant behaviour by Scheidler et al., Depickére et al.)
- varying #ants, #timestamps, #trajectories per ant, probability of making a mistake
Experiments

Synthetic dataset:

- “realistic” trajectories (follows models of ant behaviour by Scheidler et al., Depickére et al.)
- varying #ants, #timestamps, #trajectories per ant, probability of making a mistake

for real data:

\[ P(\text{error}) = 0.02 \]
Experiments

Synthetic dataset:

- “realistic” trajectories (follows models of ant behaviour by Scheidler et al., Depickére et al.)
- varying #ants, #timestamps, #trajectories per ant, probability of making a mistake

\[ \text{ALG} \quad \frac{\text{optimal fractional}}{\text{optimal fractional}} \]

![Graph showing the percentage of the optimal fractional solution against the number of ants, with lines for ILP, Greedy, and LP+Rounding.]
Conclusions

- a new approach for identifying insect trajectories
- a framework for computing consensus curves
- a new variant of the edge-disjoint paths problem
Conclusions

- a new approach for identifying insect trajectories
- a framework for computing consensus curves
- a new variant of the edge-disjoint paths problem

Future Directions

- Theoretical: approximation algorithms for SCP
Conclusions

- a new approach for identifying insect trajectories
- a framework for computing consensus curves
- a new variant of the edge-disjoint paths problem

Future Directions

- Theoretical: approximation algorithms for SCP
- Practical: computing consensus from smaller pieces
Conclusions

- a new approach for identifying insect trajectories
- a framework for computing consensus curves
- a new variant of the edge-disjoint paths problem

Future Directions

- Theoretical: approximation algorithms for SCP
- Practical: computing consensus from smaller pieces