Derivative Delay Embedding:
Online Modeling of Streaming Time Series

Zhifei Zhang (PhD student),
Yang Song, Wei Wang, and Hairong Qi

Department of Electrical Engineering & Computer Science
Outline

1. Challenges of Online Modeling
2. Derivative Delay Embedding (DDE)
3. Markov Geographic Model (MGM)
4. Experimental Results
Challenges of Online Modeling

Most modeling methods require pre-processing or assumptions:

• Segmentation
• Alignment
• Normalization

However, for the online scenario:

• Infinite time series
• Real-time
• Misalignment

Pre-processing and unrealistic assumptions are not allowed, thus misalignment challenges the online modeling
Misalignment mainly refers to the variation in phase and repeat rate of streaming time series.
The Proposed Approach

Streaming time series

Derivative Delay Embedding (DDE)

Markov Geographic Model (MGM)

- Misaligned
- Non-periodic

- Invariant to misalignment
- Real-time
- Incremental manner

- Online modeling
- Online testing
Delay Embedding (DE)

Reconstruct a latent dynamical system which generates the time series regardless of misalignment.

\[ \Phi(x_t; s, d) = (y_t, y_{t+s}, \ldots, y_{t+(d-1)s}) \]

\( \Phi \) --- estimate of \( x \)

\( s \) --- delay step

\( d \) --- embedding dimension

\( x_t \) --- state of the latent dynamical system at the time \( t \)

\( y_t \) --- observation (time series) at the time \( t \)
A Toy Examples of Delay Embedding

\[ \Phi(x_t; s, d = 2) = (y_t, y_{t+s}) = (f(t), f(t+s)) \]

- \( \Phi \) --- estimate of \( x \)
- \( s \) --- delay step
- \( d \) --- embedding dimension
- \( x_t \) --- state of the latent dynamical system at the time \( t \)
- \( y_t \) --- observation (time series) at the time \( t \)

The infinite time series becomes a trajectory in a bounded embedding space. It performs in real time.
Invariance to Misalignment

Misalignment in phase

Misaligned streaming time series generate the same trajectory in the embedding space
DE → Derivative Delay Embedding (DDE)

Invariant to misalignment of baseline

Original Time Series

Delay Embedding

Derivative Delay Embedding

Delay Embedding

Challenges of Online Modeling
Derivative Delay Embedding
Markov Geographic Model
Experimental Results

10/26/2016
Trajectory Modeling

In the embedding space, location of the states, and transition from one state to another carry the pattern of a trajectory.

Non-parametric model
- Probability the a state appear at certain location $P(x_t)$
- Transition probability $P(x_t | x_{t-1})$
- Discretized embedding space

$$S_{MGM}(X) = \sum_{j=1}^{t} P(x_j) \prod_{i=2}^{t} P(x_i | x_{i-1})$$

$$= S_G(X) \times S_M(X)$$
Markov Geographic Model (MGM)

Online update the transition and distribution of states.

Markov process

Geographic distribution

Embedding Space

Geographic Distribution

10/26/2016
Neighborhood Matching

\[ S_{\text{MGM}}(X) = \sum_{j=1}^{t} P(x_j) \prod_{i=2}^{t} P(x_i | x_{i-1}) \]

\[ = S_G(X) \times S_M(X) \]

Make the transition probability more robust to noise and unseen samples in testing.

\[ S_M(X) = \prod_{i=2}^{t} \frac{\sum_{\alpha \in N_r(\Phi'(x_i)), \beta \in N_r(\Phi'(x_{i-1}))} |\alpha;\beta|}{\sum_{k} \sum_{\gamma \in N_r(\Phi'(x_{i-1}))} |\Phi'(x_k);\gamma|} \]

\[ N_r(\Phi'(x_i)) \text{ --- the set of neighbors within radius } r \text{ around } \Phi'(x_i) \]
Online Modeling and Classification by DDE-MGM

Class $c$

Training stream

Buffer

DDE

Update

Query

Compare similarity

Geo. dist. Transition probability

Probability $P(x_t|x_{t-1})$

MGM 1

MGM c

MGM n

Testing stream $X$

10/26/2016
Experimental Results

Datasets:

- **UCI Character Trajectory** --- 2858 character samples of 20 classes, x and y axes were recorded.

- **MSR Action3D** --- 567 action samples of 20 classes performed by 10 subjects, human skeleton is recorded.
Experimental Results --- UCI Character Trajectory

The data is normalized and well aligned

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accu. (%)</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1NN-DTW</td>
<td>91.37</td>
<td>$3.9 \times 10^4$</td>
</tr>
<tr>
<td>SAX</td>
<td>89.96</td>
<td>128.85</td>
</tr>
<tr>
<td>HMM</td>
<td>57.89</td>
<td>$7.4 \times 10^3$</td>
</tr>
<tr>
<td>DDE-MGM</td>
<td>92.07</td>
<td>34.21</td>
</tr>
<tr>
<td>RBP</td>
<td>92.62</td>
<td>9.44</td>
</tr>
<tr>
<td>Projectron</td>
<td>92.62</td>
<td>110.26</td>
</tr>
<tr>
<td>BPAS</td>
<td>94.68</td>
<td>22.81</td>
</tr>
<tr>
<td>BOGD</td>
<td>90.02</td>
<td>15.24</td>
</tr>
<tr>
<td>NOGD</td>
<td>91.65</td>
<td>9.04</td>
</tr>
<tr>
<td>DDE-MGM</td>
<td>95.45</td>
<td>63.92</td>
</tr>
</tbody>
</table>

Online testing
Experimental Results --- MSR Action3D

The data is **not normalized** and **misaligned**

Half-vs-half validation

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accu. (%)</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1NN-DTW</td>
<td>74.73</td>
<td>$7.6 \times 10^4$</td>
</tr>
<tr>
<td>SAX</td>
<td>61.90</td>
<td>54.68</td>
</tr>
<tr>
<td>HMM</td>
<td>60.07</td>
<td>$2.1 \times 10^3$</td>
</tr>
<tr>
<td>DDE-MGM</td>
<td><strong>93.04</strong></td>
<td><strong>28.40</strong></td>
</tr>
<tr>
<td>RBP</td>
<td>23.41</td>
<td>20.23</td>
</tr>
<tr>
<td>Projectron</td>
<td>31.65</td>
<td>205.25</td>
</tr>
<tr>
<td>BPAS</td>
<td>30.36</td>
<td>12.25</td>
</tr>
<tr>
<td>BOGD</td>
<td>26.19</td>
<td>22.23</td>
</tr>
<tr>
<td>NOGD</td>
<td>29.96</td>
<td><strong>10.47</strong></td>
</tr>
<tr>
<td>DDE-MGM</td>
<td><strong>79.37</strong></td>
<td>80.38</td>
</tr>
</tbody>
</table>

Online testing

**Note:** The data is not normalized and misaligned.
An Example of Action Recognition

The joint of left wrist is plotted for three categories of actions shown in different colors:

- high arm wave
- horizontal arm wave
- hammer
Conclusion

- DDE is introduced to solve misalignment in online modeling.
- The non-parametric model MGM is proposed to model the trajectories in an online manner.
- Both modeling and classification are achieved in real time.
THANKS

Student Travel Grant

10/26/2016
Appendix: Parameter Setting of DE

\[ \Phi(x_t; s, d) = (y_t, y_{t+s}, \cdots, y_{t+(d-1)s}) \]

\( \Phi \) --- estimate of \( x \)

\( s \) --- delay step

\( d \) --- embedding dimension

\( x_t \) --- state of the latent dynamical system at \( t \)

\( y_t \) --- observation (time series) at the time \( t \)

\( d \) --- False nearest neighbor [M. Kennel et al., 1992]

\( s \) --- \[ 2\pi \times d \times s \times \frac{f}{f_s} \equiv 0 \mod \pi \] [J. A. Perea and J. Harer, 2013]