



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Modelování a analýza chování zvířat

Tomáš Šabata, Tomáš Borovička

České Vysoké Učení Technické v Praze
Fakulta informačních technologií

May 26, 2016



DATA SCIENCE LABORATORY



Presentation Outline

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- 1 Úvod
 - Motivace
 - Data a jejich předzpracování
- 2 Modelování
 - Přehled přístupů
 - Hidden Markov Model
- 3 Shuková analýza
 - Přehled přístupů
 - Vzdálenost mezi modely
 - Zvolené přístupy
- 4 Výsledky



Proč modelovat chování zvířat?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace

Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Hlavní myšlenka:

Rozdíly v chování zvířat jsou znatelné v příslušných modelech chování.

- Sledování změn v chování

- Porovnání chování jedinců



Odkud data pocházejí

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Poziční data
- Akce



Poziční data

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace

Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Chybějící hodnoty
- Šum

- Chybějící hodnoty
 - Linear interpolation

$$f_L(x) = f(x_0) + b_1(x - x_0)$$

$$b_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

- Chybějící hodnoty
 - Linear interpolation, f_L
 - Quadratic interpolation

$$f_Q(x) = f(x_0) + b_1(x - x_0) + b_2(x - x_0)(x - x_1)$$

$$b_2 = \frac{\frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_0}$$

- Chybějící hodnoty
 - Linear interpolation, f_L
 - Quadratic interpolation, f_Q
 - Cubic interpolation

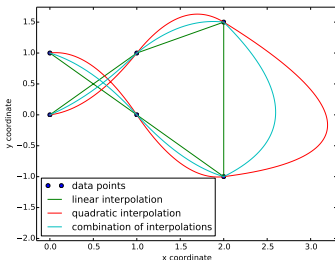
$$f_C = b_0 + b_1(x-x_0) + b_2(x-x_0)(x-x_1) + b_3(x-x_0)(x-x_1)(x-x_2)$$

$$b_3 = \frac{\frac{f(x_3)-f(x_2)}{x_3-x_2} - \frac{f(x_2)-f(x_1)}{x_2-x_1} - \frac{f(x_1)-f(x_0)}{x_1-x_0}}{x_3-x_0}$$

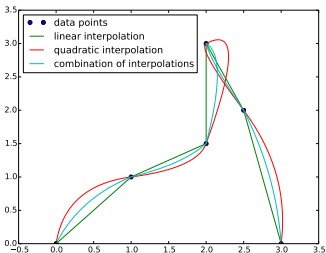
- Chybějící hodnoty
 - Linear interpolation, f_L
 - Quadratic interpolation, f_Q
 - Cubic interpolation, f_C
 - **Kombinace**

$$f(x) = \frac{f_L(x) + f_Q(x)}{2}$$

- Chybějící hodnoty
 - Linear interpolation
 - Quadratic interpolation
 - Cubic interpolation
 - **Kombinace**



- Chybějící hodnoty
 - Linear interpolation
 - Quadratic interpolation
 - Cubic interpolation
 - **Kombinace**





Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

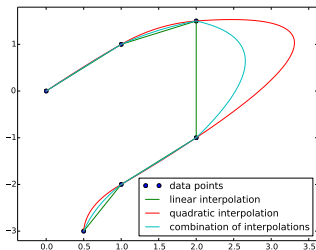
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

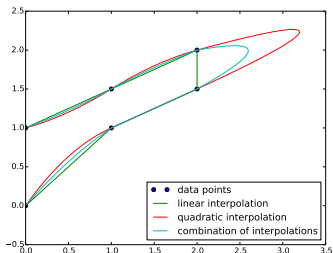
Výsledky

Reference

- Chybějící hodnoty
 - Linear interpolation
 - Quadratic interpolation
 - Cubic interpolation
 - **Kombinace**



- Chybějící hodnoty
 - Linear interpolation
 - Quadratic interpolation
 - Cubic interpolation
 - **Kombinace**





Poziční data

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Chybějící hodnoty
 - Linear interpolation
 - Quadratic interpolation
 - Cubic interpolation
 - **Kombinace**
- Filtrace šumu



Poziční data

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Chybějící hodnoty
 - Linear interpolation
 - Quadratic interpolation
 - Cubic interpolation
 - **Kombinace**
- Filtrace šumu
 - Median filter



Poziční data

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Chybějící hodnoty
 - Linear interpolation
 - Quadratic interpolation
 - Cubic interpolation
 - **Kombinace**
- Filtrace šumu
 - Median filter
 - Moving average filter



Poziční data

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Chybějící hodnoty
 - Linear interpolation
 - Quadratic interpolation
 - Cubic interpolation
 - **Kombinace**
- Filtrace šumu
 - Median filter
 - Moving average filter
 - **Kalman filter**



Kalman filter

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace

Data a jejich
předzpracování

Modelování

Přehled přístupů

Hidden Markov
Model

Shuková
analýza

Přehled přístupů

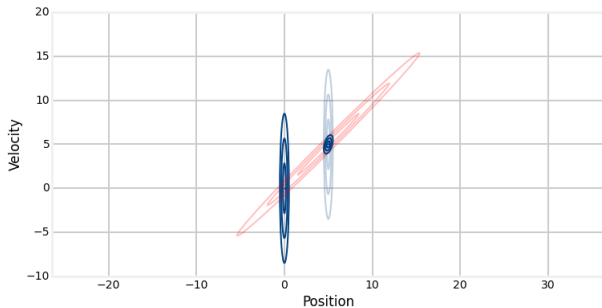
Vzdálenost mezi
modely

Zvolené přístupy

Výsledky

Reference

Proč Kalmanův filtr?



Kalman filter

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

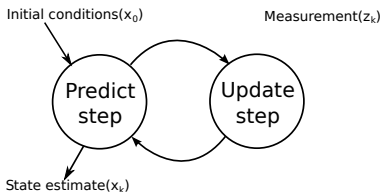
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Modelování linárního systému Skrytý markovský model se spojitými stavy



Kalman filter

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace

Data a jejich
předzpracování

Modelování

Přehled přístupů

Hidden Markov
Model

Shuková
analýza

Přehled přístupů

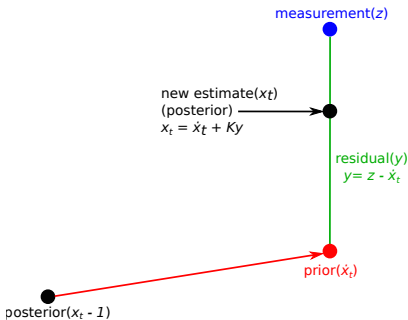
Vzdálenost mezi
modely

Zvolené přístupy

Výsledky

Reference

Co je uvnitř?



Kalman filter

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování
Přehled přístupů
Hidden Markov
Model

Shuková
analýza
Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Prediction step

$$x^- = Fx + Bu$$

$$P^- = FPF^T + Q$$

Update step

$$y = z - Hx^-$$

$$S = HP^-H^T + R$$

$$K = P^-H^T S^{-1}$$

$$x = x^- Ky$$

$$P = (I - KH)P^-$$

Legend

- x - state
- P - uncertainty covariance
- Q - process uncertainty
- u - motion vector
- B - control transition matrix
- F - state transition matrix
- H - measurement function
- R - state uncertainty
- S - measurement space
- K - kalman gain

Kalman filter

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Prediction step

$$x^- = Fx + Bu$$

Predicted (*a priori*) state estimate

$$P^- = FPF^T + Q$$

Predicted (*a priori*) estimate covariance

Update step

$$y = z - Hx^-$$

Innovation or measurement residual

$$S = HP^-H^T + R$$

Innovation (or residual) covariance

$$K = P^-H^T S^{-1}$$

Map system uncertainty into optimal kalman gain

$$x = x^- Ky$$

Updated (*a posteriori*) state estimate

$$P = (I - KH)P^-$$

Updated (*a posteriori*) estimate covariance



Presentation Outline

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- 1 Úvod
 - Motivace
 - Data a jejich předzpracování

- 2 Modelování
 - Přehled přístupů
 - Hidden Markov Model

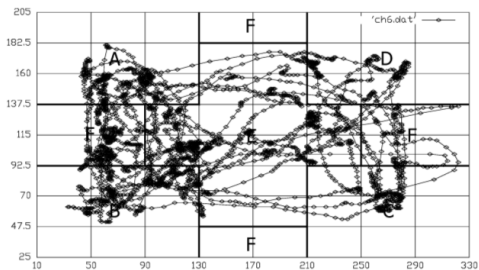
- 3 Shuková analýza
 - Přehled přístupů
 - Vzdálenost mezi modely
 - Zvolené přístupy

- 4 Výsledky



Related work

Modelování chování myši v labyrintu [1] Hidden Markov Models



```

qb1 = BEEBBFFAAEE
qb2 = EDECEE BEEFEAEDDDF
qb3 = FCCFFFAFEEBEBBEFFDD
qb4 = DDFFCBBFAEEBBEEBEEFFFC
qb5 = CCCFEEDDDDEDDDEAAFFBFBFFBBBFF
qb6 = FFAAEFBFBEB
qb7 = BEEEA
qb8 = AFFBFAAEEEBEBEE

```

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

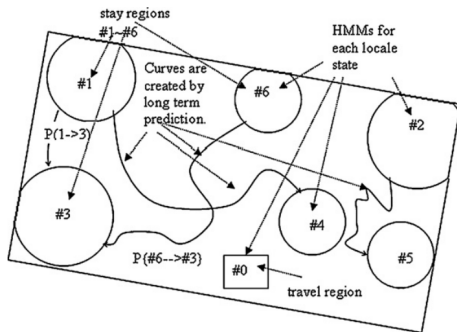
Výsledky

Reference



Related work

Modelování chování skotu [2] Hidden Markov Models



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Prehled přístupů
Hidden Markov
Model

Shuková
analýza

Prehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference



Related work

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

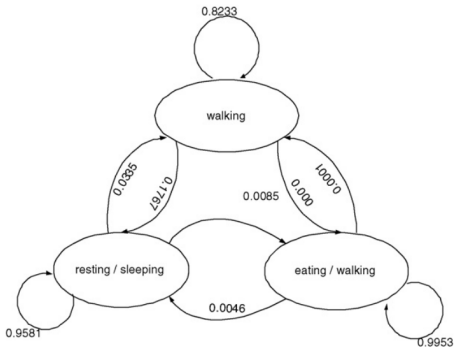
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Modelování chování skotu [2]



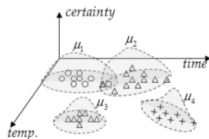
Modelování chování lidí [3] Fuzzy Q-state learning + Agglomerative fuzzy clustering

State S_1 : drinking coffee (○)
 S_2 : watching TV (△)
 S_3 : having a meal (+)

Action a_1 : serving coffee
 a_2 : turning on TV
 a_3 : serving a meal

Context attributes
time and temperature

Fuzzy partition for the state

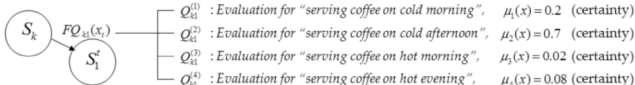


**Linguistic description
of each membership function**

- μ_1 : "cold morning"
- μ_2 : "cold afternoon"
- μ_3 : "hot morning"
- μ_4 : "hot evening"

Linguistic description for fuzzy-state transition

$x_i = (15, 5)$: "time = 15:00, temp. = 5°C"



Přístupy modelování sekvencí

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

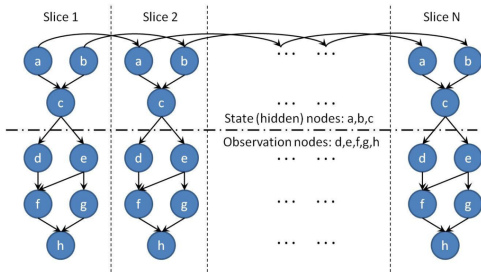
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

■ Dynamic Bayesian network



Přístupy modelování sekvencí

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

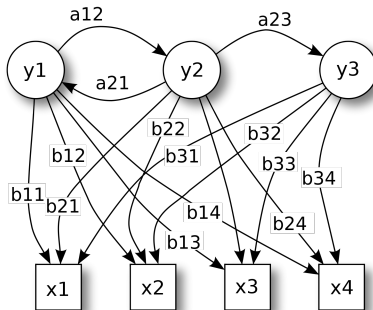
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Dynamic Bayesian network
- Hidden Markov Models



Přístupy modelování sekvencí

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

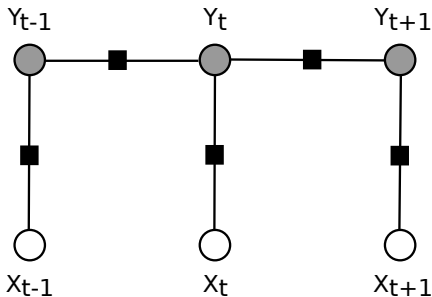
Modelování
Přehled přístupů
Hidden Markov
Model

Shuková
analýza
Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Dynamic Bayesian network
- Hidden Markov Models
- Linear chain Conditional Random fields



Přístupy modelování sekvencí

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

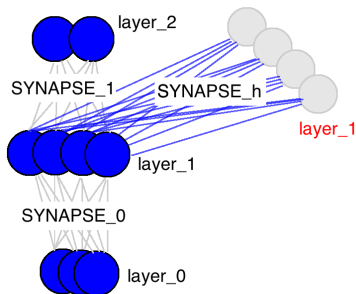
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Dynamic Bayesian network
- Hidden Markov Models
- Conditional Random fields
- LSTM RNN



Co je Hidden Markov Model (HMM)?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

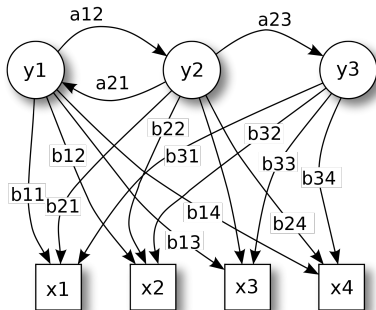
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$$\lambda = (A, B, \pi)$$





Co je Hidden Markov Model (HMM)?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
**Hidden Markov
Model**

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$$\lambda = (A, B, \pi)$$

Co je Hidden Markov Model (HMM)?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
**Hidden Markov
Model**

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$$\lambda = (A, B, \pi)$$

- Množina skrytých stavů $Y = \{y_1, y_2, \dots, y_N\}$ a pozorovaných výstupů $X = \{x_1, x_2, \dots, x_M\}$

Co je Hidden Markov Model (HMM)?



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$$\lambda = (A, B, \pi)$$

- Množina skrytých stavů $Y = \{y_1, y_2, \dots, y_N\}$ a pozorovaných výstupů $X = \{x_1, x_2, \dots, x_M\}$
- Sekvence stavů $Q = q_1 q_2 q_3 \dots q_T$ a sekvence výstupů $O = o_1 o_2 o_3 \dots o_T$

Co je Hidden Markov Model (HMM)?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$$\lambda = (A, B, \pi)$$

- Množina skrytých stavů $Y = \{y_1, y_2, \dots, y_N\}$ a pozorovaných výstupů $X = \{x_1, x_2, \dots, x_M\}$
- Sekvence stavů $Q = q_1 q_2 q_3 \dots q_T$ a sekvence výstupů $O = o_1 o_2 o_3 \dots o_T$
- Pravděpodobnosti přechodů (transitions) $A = \{a_{ij}\}$
 $a_{ij} = P(q_t = y_j | q_{t-1} = y_i), \quad 1 \leq i, j \leq N$

Co je Hidden Markov Model (HMM)?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$$\lambda = (A, B, \pi)$$

- Množina skrytých stavů $Y = \{y_1, y_2, \dots, y_N\}$ a pozorovaných výstupů $X = \{x_1, x_2, \dots, x_M\}$
- Sekvence stavů $Q = q_1 q_2 q_3 \dots q_T$ a sekvence výstupů $O = o_1 o_2 o_3 \dots o_T$
- Pravděpodobnosti přechodů (transitions) $A = \{a_{ij}\}$
 $a_{ij} = P(q_t = y_j | q_{t-1} = y_i), \quad 1 \leq i, j \leq N$
- Pravděpodobnosti pozorovaných symbolů (emissions)
 $B = \{b_{i,j}\}$
 $b_{i,j} = P(o_t = x_j | q_t = y_i), \quad 1 \leq i \leq N, \quad 1 \leq j \leq M$

Co je Hidden Markov Model (HMM)?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$$\lambda = (A, B, \pi)$$

- Množina skrytých stavů $Y = \{y_1, y_2, \dots, y_N\}$ a pozorovaných výstupů $X = \{x_1, x_2, \dots, x_M\}$
- Sekvence stavů $Q = q_1 q_2 q_3 \dots q_T$ a sekvence výstupů $O = o_1 o_2 o_3 \dots o_T$
- Pravděpodobnosti přechodů (transitions) $A = \{a_{ij}\}$
 $a_{ij} = P(q_t = y_j | q_{t-1} = y_i), \quad 1 \leq i, j \leq N$
- Pravděpodobnosti pozorovaných symbolů (emissions)
 $B = \{b_{i,j}\}$
 $b_{i,j} = P(o_t = x_j | q_t = y_i), \quad 1 \leq i \leq N, \quad 1 \leq j \leq M$
- Pravděpodobnosti počátečních stavů $\pi = \{\pi_i\}$
 $\pi_i = P(q_1 = y_i), \quad 1 \leq i \leq N$

Jak HMM použít?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
**Hidden Markov
Model**

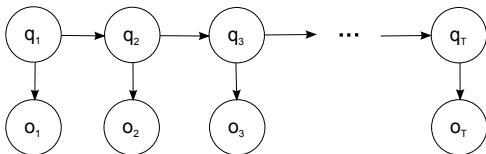
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$P(O, Q)$ - Trellis diagram



Jak HMM použít?



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

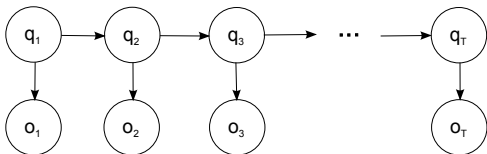
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$P(O, Q)$ - Trellis diagram



$$P(o_1, \dots, o_T, q_1, \dots, q_T) = P(q_1)P(o_1|q_1) \prod_{k=2}^T P(q_k|q_{k-1})P(o_k|q_k)$$

$$P(o_1, \dots, o_T, q_1, \dots, q_T) = \pi_{q_1} b_{q_1, o_1} \prod_{k=2}^T a_{q_{k-1}, q_k} b_{q_k, o_k}$$

Jak HMM použít?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

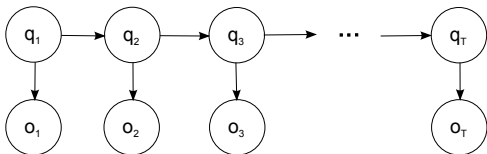
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$P(O, Q)$ - Trellis diagram



$$P(o_1, \dots, o_T, q_1, \dots, q_T) = P(q_1)P(o_1|q_1) \prod_{k=2}^T P(q_k|q_{k-1})P(o_k|q_k)$$

$$P(o_1, \dots, o_T, q_1, \dots, q_T) = \pi_{q_1} b_{q_1, o_1} \prod_{k=2}^T a_{q_{k-1}, q_k} b_{q_k, o_k}$$

- Vybavení modelu? Viterbiho algoritmus - $\arg \max_Q P(O, Q|\lambda)$

Jak HMM použít?

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

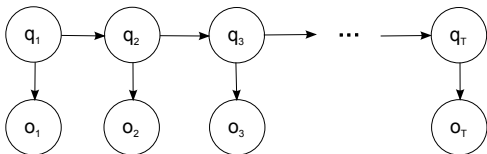
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

$P(O, Q)$ - Trellis diagram



$$P(o_1, \dots, o_T, q_1, \dots, q_T) = P(q_1)P(o_1|q_1) \prod_{k=2}^T P(q_k|q_{k-1})P(o_k|q_k)$$

$$P(o_1, \dots, o_T, q_1, \dots, q_T) = \pi_{q_1} b_{q_1, o_1} \prod_{k=2}^T a_{q_{k-1}, q_k} b_{q_k, o_k}$$

- Vybavení modelu? Viterbiho algoritmus - $\arg \max_Q P(O, Q|\lambda)$
- Učení modelu? EM algoritmus - $\arg \max_\lambda P(O, Q|\lambda)$



Presentation Outline

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- 1 Úvod
 - Motivace
 - Data a jejich předzpracování
- 2 Modelování
 - Přehled přístupů
 - Hidden Markov Model
- 3 Shuková analýza
 - Přehled přístupů
 - Vzdálenost mezi modely
 - Zvolené přístupy
- 4 Výsledky



Existující algoritmy shlukové analýzy HMM

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Spektrální shlukování pomocí Bhattacharya divergence[4]
- Spektrální shlukování pomocí Kullback-Leiblerovy divergence[5]
- Variational hierarchical EM [6]



Existující algoritmy shlukové analýzy HMM

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Spektrální shlukování pomocí Bhattacharya divergence[4]
- Spektrální shlukování pomocí Kullback-Leiblerovy divergence[5]
- Variational hierarchical EM [6]

V čem je problém?



Existující algoritmy shlukové analýzy HMM

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Spektrální shlukování pomocí Bhattacharya divergence[4]
- Spektrální shlukování pomocí Kullback-Leiblerovy divergence[5]
- Variational hierarchical EM [6]

V čem je problém? → Normálně rozdělené emissions!



Existující algoritmy shlukové analýzy HMM

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Spektrální shlukování pomocí Bhattacharya divergence[4]
- Spektrální shlukování pomocí Kullback-Leiblerovy divergence[5]
- Variational hierarchical EM [6]

V čem je problém? → Normálně rozdělené emissions!
Jak z toho ven?



Existující algoritmy shlukové analýzy HMM

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Spektrální shlukování pomocí Bhattacharya divergence[4]
- Spektrální shlukování pomocí Kullback-Leiblerovy divergence[5]
- Variational hierarchical EM [6]

V čem je problém? → Normálně rozdělené emissions!
Jak z toho ven? → definovat vzdálenost jinak.

Euklidovská vzdálenost



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování
Přehled přístupů
Hidden Markov
Model

Shuková
analýza
Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Euklidovská vzdálenost[7]

- Euklidovská vzdálenost mezi řádky matice B

$$d_{ec}(\lambda, \lambda') = \sqrt{\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^M \|b_{ik} - b'_{ik}\|^2}$$

- Minimalizovaná Euklidovská vzdálenost

$$d_{mec}(\lambda, \lambda') = \sqrt{\frac{1}{N} \sum_{i=1}^N \min_j \sum_{k=1}^M \|b_{ik} - b'_{jk}\|^2}$$

Pravděpodobnostní rozdělení B je z hlediska podobnosti dvou HMM nejvýznamější [8].

Statistická vzdálenost I



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Kullback-Leiblerova divergence[7]
výpočetně náročné \rightarrow aproximace.

$$1 \quad d_{KL}(\lambda, \lambda') = \int_O \frac{1}{G(O)} \log \frac{P(O|\lambda)}{P(O|\lambda')} P(O|\lambda) dO$$

2 předpoklad: $Q = Q' = Q_{\text{opt}}$, modely si jsou podobné

$$3 \quad d_{\text{vit}}(\lambda, \lambda') = \int_O \frac{1}{G(O)} \log \frac{P(Q_{\text{opt}}, O|\lambda)}{P(Q_{\text{opt}}, O|\lambda')} P(O|\lambda) dO$$

4 předpoklad: Markovský řetězec je ergodický

$$5 \quad d_{\text{vit}}(\lambda, \lambda') = \frac{1}{G(O)} \log \frac{P(Q_I, O|\lambda)}{P(Q_I, O|\lambda')} P(O|\lambda) + \epsilon$$



6

$$d_{Vit}(\lambda, \lambda') - \epsilon = \frac{1}{G(O)} \sum_{t=1}^{T-1} (\log a_{y_t, y_{t+1}} - \log a'_{y_t, y_{t+1}}) + \frac{1}{G(O)} \sum_{t=1}^T (\log b_{y_t, x_t} - \log b'_{y_t, x_t})$$

7 Délka sekvence O je dostatečně dlouhá

8

$$d_{Vit}(\lambda, \lambda') \approx \tilde{\delta}_{Vit} = \sum_{i,j} a_{ij} \pi_i (\log a_{ij} - \log a'_{ij}) + \sum_{i,k} b_{ik} \pi_i (\log b_{ij} - \log b'_{ij})$$



1 Hierarchistické shlukování

- Linkage ze vzdálenostní matice (single, average, complete)

2 Spektrální shlukování

- podle Shi a Malik (2000)
- podle Ng, Jordan a Weiss (2002)
- Vytvoříme matici vzdáleností X
- Reprezentace této pomocí podobnostního grafu
- Similarity matrix W , $w_{ij} = e^{-\frac{\|x_{ij}\|^2}{2\sigma^2}}$
- Degree matrix D , $D_{i,i} = \sum_{j \in V} w_{ij}$, $D_{i,j;i \neq j} = 0$

Spektrální shlukování - Shi and Malik

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

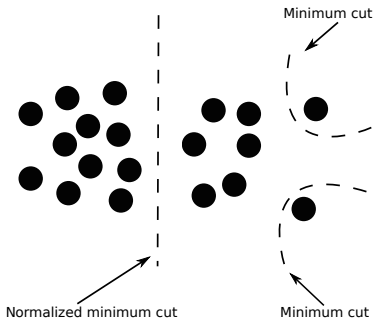
Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- 1 Výpočet vlastních čísel $(D - W)y = \lambda Dy$
- 2 Pomocí druhého nejmenšího vlastního čísla rekurzivně dělíme graf.





Presentation Outline

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- 1 Úvod
 - Motivace
 - Data a jejich předzpracování
- 2 Modelování
 - Přehled přístupů
 - Hidden Markov Model
- 3 Shuková analýza
 - Přehled přístupů
 - Vzdálenost mezi modely
 - Zvolené přístupy
- 4 Výsledky

Evaluace modelu



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- Vizuální evaluace doménovým expertem
- Evaluace pomocí Viterbiho sekvence

Label	Acc. Mean	Acc. Variance
S1	93.45%	2.22
S2	99.76%	0.69
S3	93.94%	2.82
S4	94.83%	4.12
S5	85.73%	3.93
Přesnost	83.86%	4.76

Výsledky shlukování

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

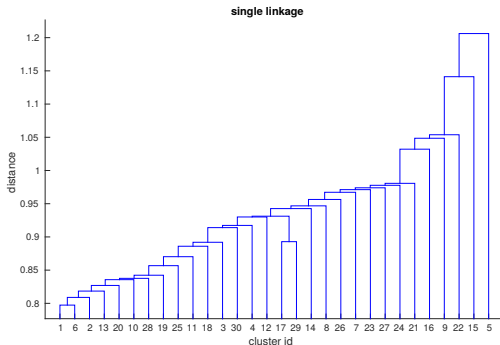
Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Hierarchistické shlukování

■ single linkage





Výsledky shlukování

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

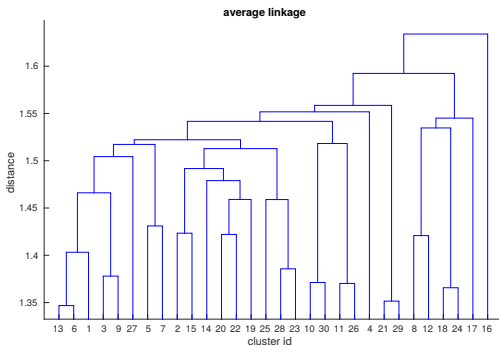
Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

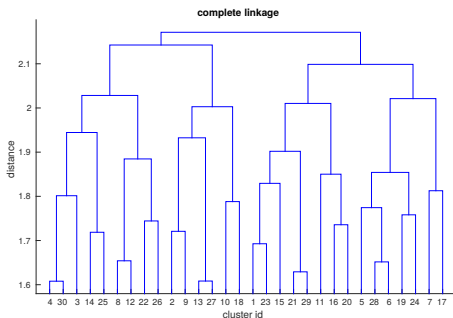
Hierarchistické shlukování

- Single linkage
- Average linkage



Hierarchistické shlukování

- Single linkage
- Average linkage
- Complete linkage





Výsledky shlukování

Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

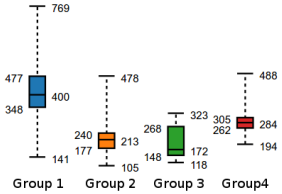
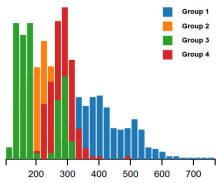
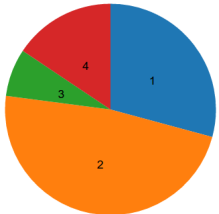
Úvod
Motivace
Data a jejich
předzpracování

Modelování
Přehled přístupů
Hidden Markov
Model

Shuková
analýza
Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky
Reference

Spektrální shlukování



Reference I



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- 1 MATETIĆ, Maja; RIBARIĆ, Slobodan; IPŠIĆ, Ivo. Qualitative modelling and analysis of animal behaviour. Applied Intelligence, 2004, 21.1: 25-44.
- 2 Y. Guo, G. Poulton, P. Corke, G.J. Bishop-Hurley, T. Wark, D.L. Swain, Using accelerometer, high sample rate GPS and magnetometer data to develop a cattle movement and behaviour model, Ecological Modelling, Volume 220, Issue 17, 10 September 2009, Pages 2068-2075, ISSN 0304-3800
- 3 LEE, Sang Wan; KIM, Yong Soo; BIEN, Zeungnam. A nonsupervised learning framework of human behavior patterns based on sequential actions. Knowledge and Data Engineering, IEEE Transactions on, 2010, 22.4: 479-492.
- 4 Tony Jebara, Yingbo Song, and Kapil Thadani. Spectral clustering and embedding with hidden Markov models. In Machine Learning: ECML 2007, pages 164–175. 2007

Reference II



Modelování a
analýza
chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková
analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

- 5 Shi Zhong and Joydeep Ghosh. A unified framework for model-based clustering. *The Journal of Machine Learning Research*, 4:1001–1037, 2003
- 6 COVIELLO, Emanuele; CHAN, Antoni B.; LANCKRIET, Gert RG. Clustering hidden Markov models with variational HEM. *The Journal of Machine Learning Research*, 2014, 15.1: 697-747.
- 7 FALKHAUSEN, Markus; REININGER, Herbert; WOLF, Dietrich. Calculation of distance measures between hidden Markov models. In: *EUROSPEECH*. 1995.
- 8 JUANG, Biing-Hwang Fred; RABINER, Lawrence R. A probabilistic distance measure for hidden Markov models. *AT&T technical journal*, 1985, 64.2: 391-408.



Modelování a analýza chování zvířat

Tomáš
Šabata,
Tomáš
Borovička

Úvod

Motivace
Data a jejich
předzpracování

Modelování

Přehled přístupů
Hidden Markov
Model

Shuková analýza

Přehled přístupů
Vzdálenost mezi
modely
Zvolené přístupy

Výsledky

Reference

Děkujeme za pozornost.

- tomas.sabata@fit.cvut.cz
- tomas.borovicka@fit.cvut.cz