

Klastrowanie sekwencji zdarzeń

Integracja Systemów

Założenia

System będzie

- Obierał i zachowywał dane odnośnie zdarzeń (w formie batchowej)
- Odbierał dane odnośnie znalezionych częstych sekwencji
- Po odebraniu obu powyższych zdarzeń będzie klastrował dane na podstawie różnic znalezionych przy najczeststch sekwencjach
- Znalezione klastry będzie wysyłał do zainteresowanych komponentów

Technologie



Architektura

System będzie składał się komponentów:

- Webserwis odpowiedzialny za obieranie, wysyłanie oraz prasowanie danych
- Podsystem odpowiedzialny za wytworzenie reguł oraz ich ekstrakcje z danych
- Postem klastrujący dane

Komunikaty

System będzie obierał 2 rodzaje komunikatów:

- Dane odnośnie zdarzeń (zdarzenia)
- Dane odnośnie sekwencji (sekwencje)

System będzie wysyłał dane odnośnie znalezionych klastrow (klastry)

Szkic komunikatów: zdarzenia

Szkic przykładowego komunikatu:

```
{ events: [  
  [ //time window  
    { time: 12345 //timestamp  
      what: String //event kind eg. turn left  
      data: String // additional data eg. turn degree  
      who: 123 // long id of event creator  
      id: 1234 //long id of event  
    },  
    {...}  
  ],  
  [...]  
]}
```


Szkic komunikatów: sekwencje

Szkic przykładowego komunikatu:

```
{sequences:[  
  { //single sequence  
    id: Long,  
    order: [  
      String, //event kinds in order in time window  
      String,  
      String  
    ]  
  },  
  [...]  
]}
```

Szkic komunikatów: klastry

Szkic przykładowego komunikatu:

```
{clusters:[
  {sequence: {
    id: Long //if of sequence
    order: [
      what, //event kinds in order in time window
      what,
      what
    ]
  }
  groups: [
    [
      id, //id of event
      id,
    ],
    [...]
  ]
},
{...}
]}
```