

ICDE2016 & WWW2016 勉強会

**Research Session 8A:
Data Explorations and Event Analytics**

担当：杉浦 健人（名大）

Event Regularity and Irregularity in a Time Unit

Lijian Wan and Tingjian Ge (Univ. of Massachusetts)

目的 あるTime Unit毎に繰り返されるイベント系列から
規則性を導き，異常検知に利用する

例 飛行機のフライトにおける規則性 **離陸後降着装置をしまう**

■ $s_1 = \langle \text{stop, takeOff, retractLandingGear, ... , stop} \rangle$

■ $s_2 = \langle \text{stop, takeOff, retractLandingGear, ... , stop} \rangle$

■ $s_3 = \langle \text{stop, takeOff, retractLandingGear, ... , stop} \rangle$

フライト時の異常検知 **降着装置をしまっていない!**

■ $s_{new} = \langle \text{stop, takeOff, clime, cruise, ...} \rangle$

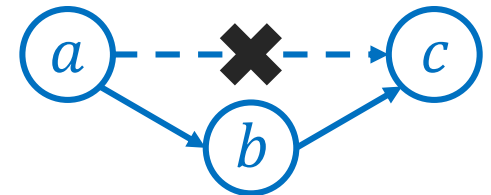
規則性の導出 | Regularity Model

規則的なイベントの組：全系列 S の中で τ 割以上出現

▶ 規則的なイベントの組をグラフで表す

▶ グラフ構築のアイデア：メッセージの伝播

例 $S = \{\langle a, b, c \rangle, \langle a, b, c \rangle, \langle a, b, c \rangle\}, \tau = 0.9$



1. 隣のイベントにメッセージを送る： $a \rightarrow b, b \rightarrow c$

■ 規則的ならグラフに辺を加える

2. 受け取ったメッセージを隣に渡す： $a(\rightarrow b) \rightarrow c$

■ 規則的かつ到達可能な経路がなければ辺を加える

3. 伝播が終わるまで繰り返す

異常検知 | Regularity Score

規則的かどうかのスコアを“ping”のアイデアを基に計算

▶ メッセージを送って応答がなければ異常

例 $s = \langle a, b, a \rangle$, 右図のRegularity Model 

1. Regularity Scoreを初期化

■ $\{\{0,0\}, \{0,0\}, \{0,0\}\}$: i 番目のスコア $\{pre[i], post[i]\}$

2. 隣のイベントにメッセージを送る : $a \rightarrow b, b \rightarrow a$

■ グラフ中に辺があれば対応するスコアを 1 に
 $\{\{0,1\}, \{1,0\}, \{0,0\}\}$:

3. メッセージを伝播して繰り返し, スコアの平均をとる

イベントの概念階層と最適化

論文中ではイベントの概念階層も考慮

- ▶ 具体的なイベントほどRegularity Scoreが高い

計算量を減らすための工夫も紹介

- ▶ 伝播によって辺が追加される確率を計算し、確率がしきい値以下になったら伝播を終了
 - ▶ precision = 1 かつ recall = 0.99 で高速化に成功
- ▶ イベントの組を含む系列の数を数える際に、Blended Bitmap Setを使って近似
 - ▶ メモリ使用量も 1/6 程度に