GRETA: Graph-based Real-time Event Trend Aggregation

Olga Poppe[1], **Chuan Lei**[2],
Elke A. Rundensteiner[3], and David Maier[4]

[1]Microsoft Gray Systems Lab, [2]IBM Research AI,
[3]Worcester Polytechnic Institute, [4]Portland State University

VLDB, August 29, 2018

Supported by NSF grants IIS 1018443, IIS 1343620, IIS 1560229, CRI 1305258

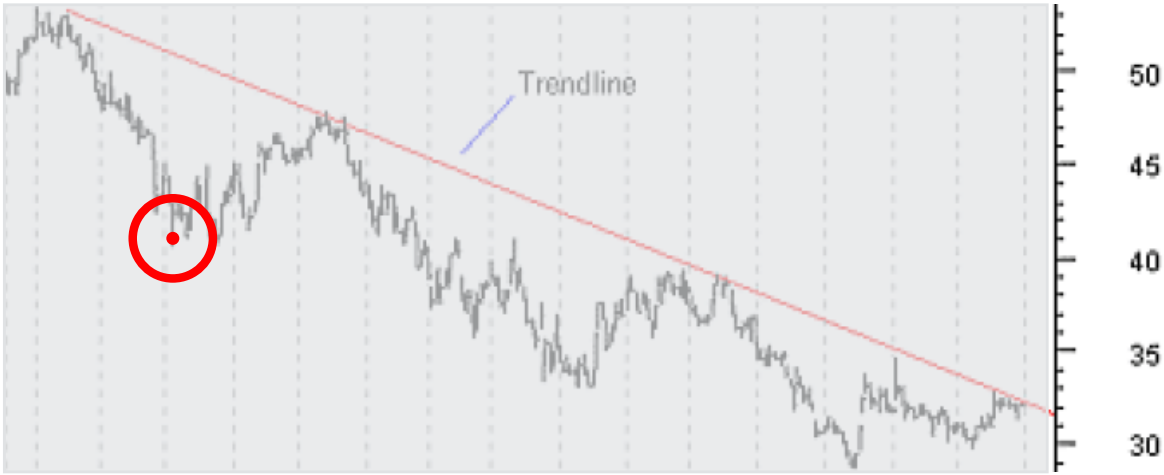
Motivation – Algorithmic Trading

Goal: Reliable actionable insights about the stream

Solution: Each event is considered in the context of other events in the stream

Picture source: http://www.businessxack.com/how-to-know-the-stock-market-trend/1303

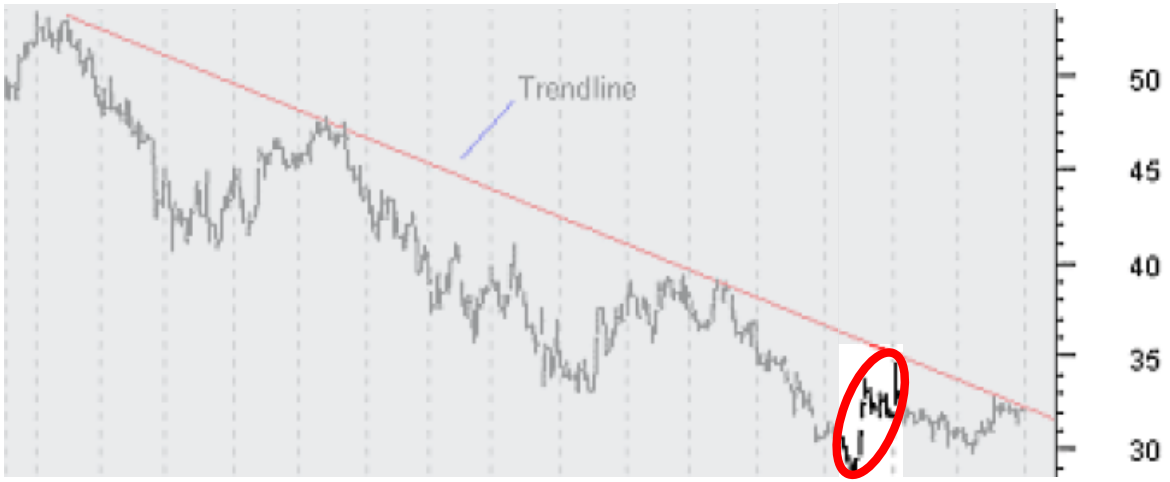
Algorithmic Trading



- **Single event** = Single stock value
- **Event sequence** = Stock down trend of fixed length
- **Event trend** = Stock down trend of any length

Picture source: http://www.businessxack.com/how-to-know-the-stock-market-trend/1303

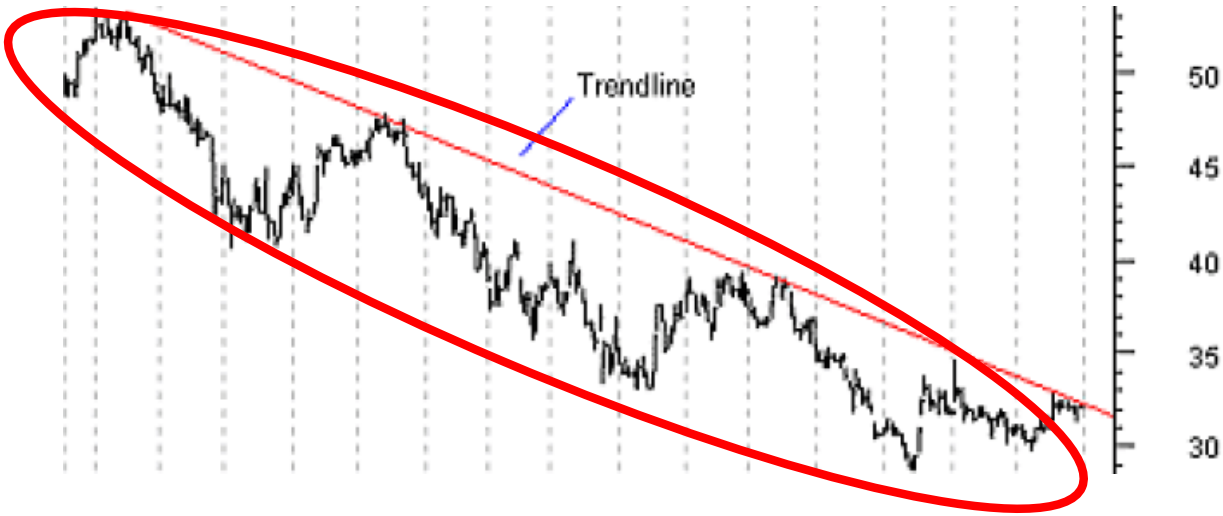
Algorithmic Trading



- **Single event** = Single stock value
- **Event sequence** = Stock down trend of fixed length
- **Event trend** = Stock down trend of any length

Picture source: http://www.businessxack.com/how-to-know-the-stock-market-trend/1303

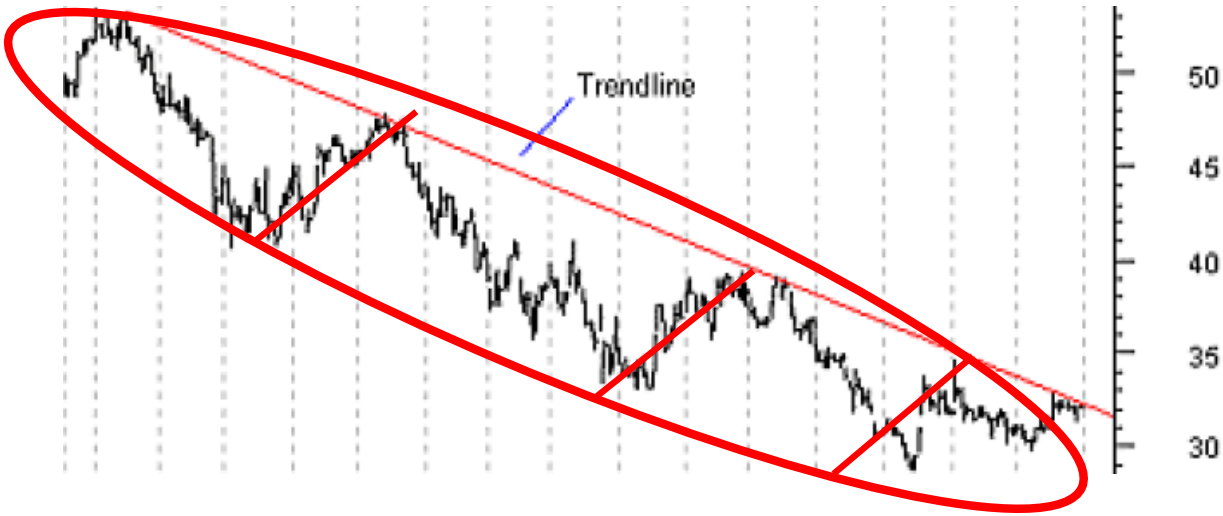
Algorithmic Trading



- **Single event** = Single stock value
- **Event sequence** = Stock down trend of fixed length
- **Event trend** = Stock down trend of any length

Picture source: http://www.businessxack.com/how-to-know-the-stock-market-trend/1303

Algorithmic Trading



- **Single event** = Single stock value
- **Event sequence** = Stock down trend of fixed length
- **Event trend** = Stock down trend of any length under the **skip-till-next-match** semantics*

* E.Wu, Y.Diao, and S.Rizvi. High-performance Complex Event Processing over streams. SIGMOD, pages 407-418, 2006.

Event Trends in Other Streaming Applications

Traffic control

Event trend:
Aggressive driving

Health care

Event trend:
Irregular heart rate

Cluster monitoring

Event trend:
Uneven load distribution

E-commerce

Event trend:
Items often bought together

Stock market

Event trend:
Head-and-shoulders

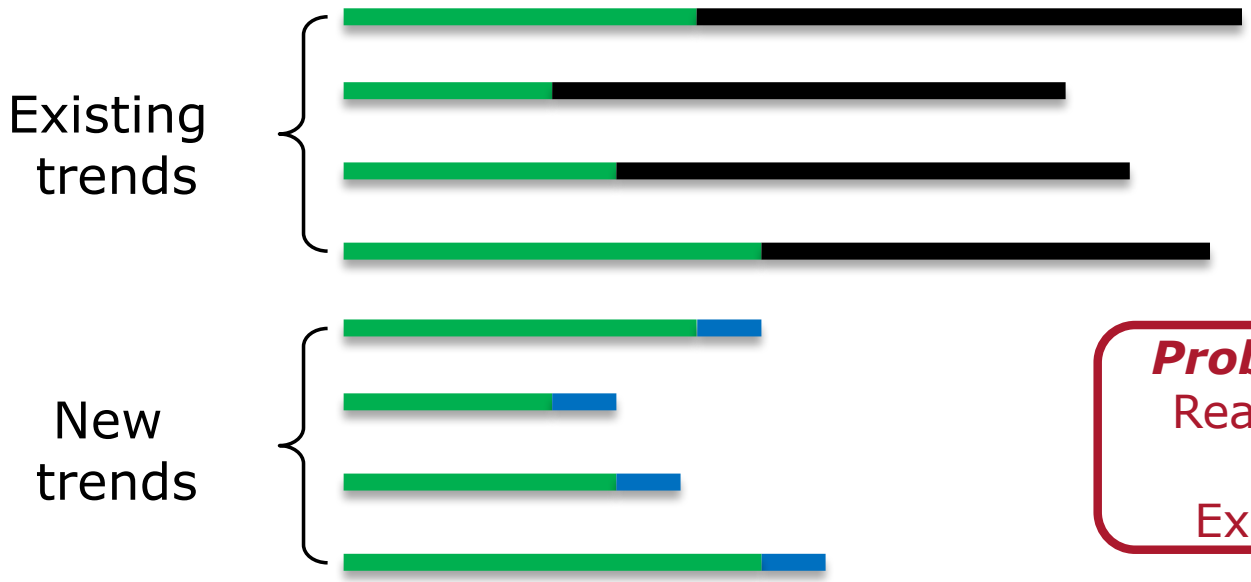
Financial fraud

Event trend:
Circular check kite

Complexity of Event Trend Analytics

Complexity of Event Trend Analytics

Complexity of Event Trend Analytics



Problem Statement
Real-time Response despite Exponential Costs

- Exponential number of trends
- Arbitrary length of a trend
- Complex event inter-dependencies in a trend

=> **Exponential time complexity**

Existing Two-Step Approachs

Step 2: **Event Trend Aggregation**

Step 1:
Event Trend Construction

Exponential
time & space complexity

Event Trend Aggregation Query

RETURN sector, COUNT(*)
PATTERN Stock S+
WHERE [company] **AND** S.price > **NEXT**(S).price
GROUP-BY sector **WITHIN** 30 min **SLIDE** 1 min

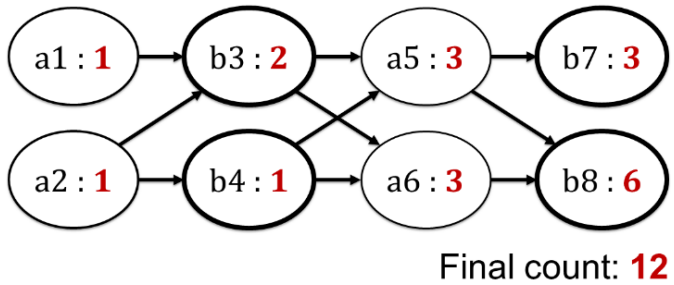
Event Stream

Transaction event
- Sector id
- Company id
- Price
- Time

Picture source: http://www.zerohedge.com/news/2015-12-05/dozens-global-stock-markets-are-already-crashing-not-seen-numbers-these-2008

GRETA: Graph-based Real-time Event Trend Aggregation

Graph-Based Event Trend Aggregation



Quadratic time & **linear** space complexity

Event Trend Aggregation Query

RETURN sector, COUNT(*)
PATTERN Stock S+
WHERE [company] **AND** S.price > **NEXT**(S).price
GROUP-BY sector **WITHIN** 30 min **SLIDE** 1 min

Event Stream

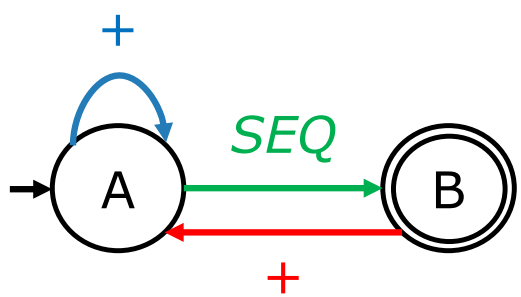
Transaction event
- Sector id
- Company id
- Price
- Time

Graph Template

Nested Kleene Pattern

$P = (SEQ(A+, B))+$

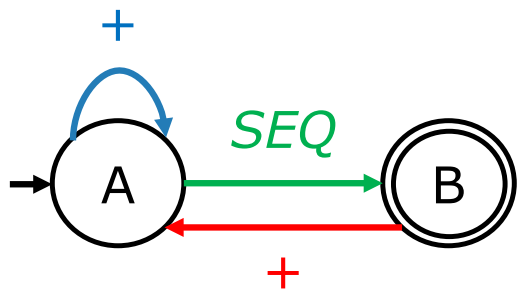
a's are preceded by a's and b's

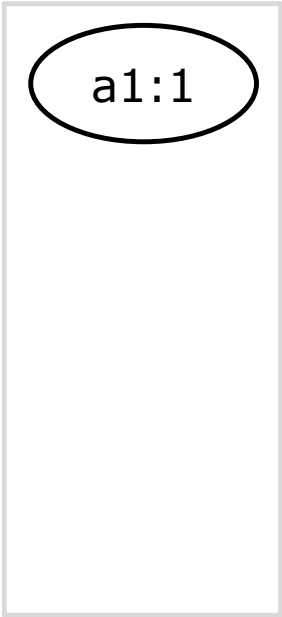


b's are preceded by a's

Start type **End type**

Graph-Based Trend Aggregation

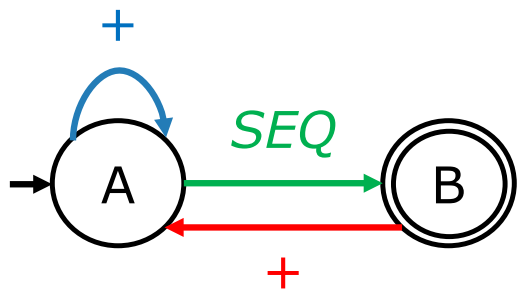


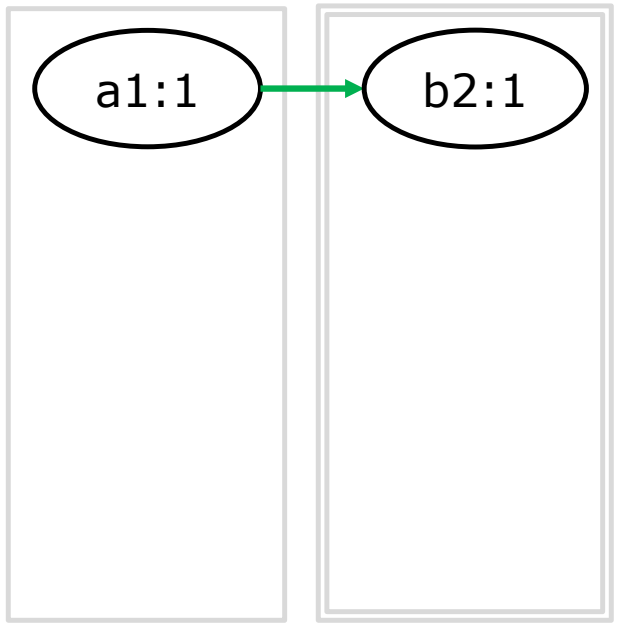


Event trends:
(a1,

Final count: 0

Graph-Based Trend Aggregation

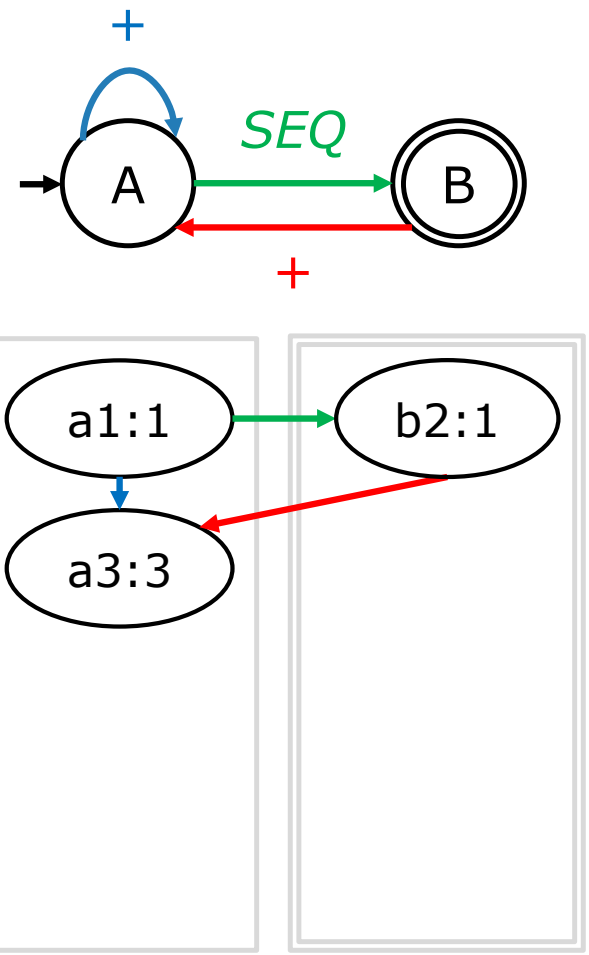




Event trends:
(a1,b2)

Final count: 1

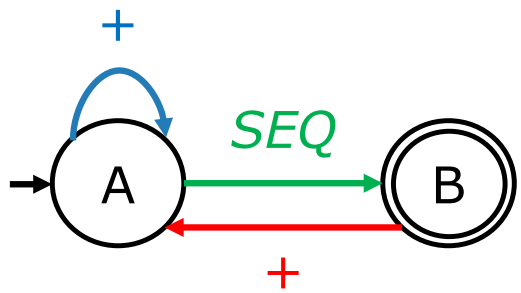
Graph-Based Trend Aggregation

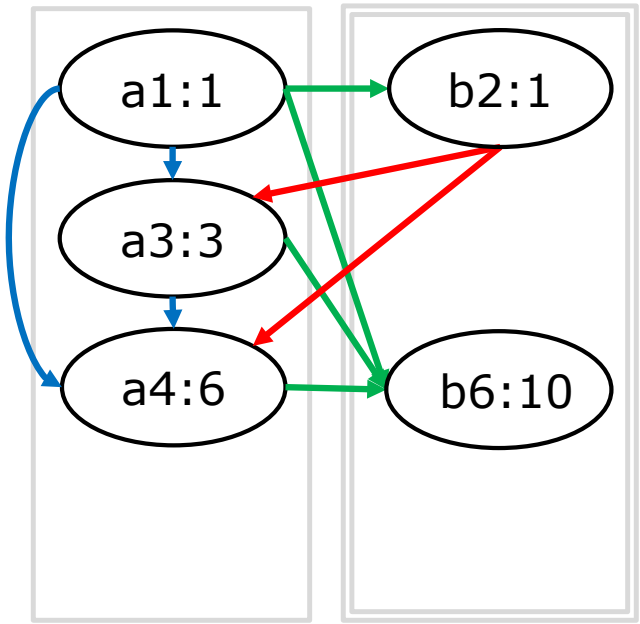


Event trends:
(a1,b2,a3,
(a1,a3,
(a3,

Final count: 1

Graph-Based Trend Aggregation



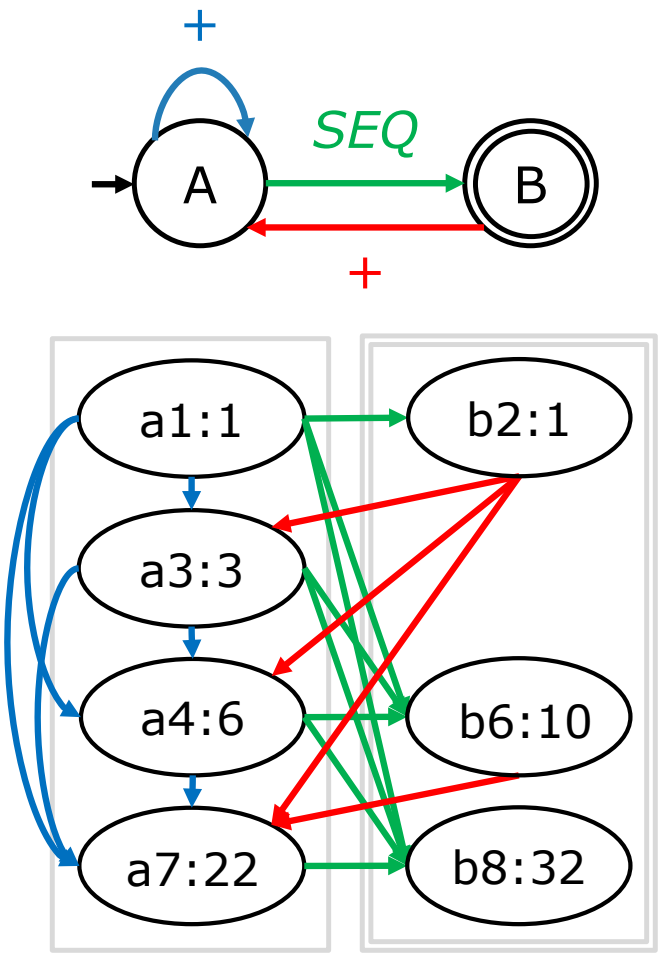


Event trends:
(a1,b6),…
(a1,a3,b6),…
(a1,b2,a3,b6),…
(a1,b2,a3,a4,b6)

Final count: 11

Graph-Based Trend Aggregation

Our GRETA approach



Quadratic time & **linear** space complexity

Final count: 43

Existing two-step approaches

Event trends:
(a1,b8),…
(a1,a3,b8),…
(a1,b6,a7,b8),…
(a1,a3,a4,a7,b8),…
(a1,b2,a3,a4,a7,b8),…

Exponential time & space complexity

Experimental Setup

Execution infrastructure:

Java 7, 1 Linux machine with 16-core

3.4 GHz CPU and 128GB of RAM

Data sets:

- **ST**: Stock real data set

  Event trends = Stock market trends

- **LR**: Linear road benchmark data set

  Event trends = Vehicle trajectories

- **CL**: Cluster monitoring synthetic data set

  Event trends = Load distribution trends

ST: Stock trade traces. http://davis.wpi.edu/datasets/Stock Trace Data/
LR: A.Arasu, M.Cherniack, E.Galvez, D.Maier, A.S.Maskey, E.Ryvkina, M.Stonebraker, and R.Tibbetts. Linear road: A stream data management benchmark. In VLDB, pages 480-491, 2004.

Event Aggregation Approaches

Existing two-step approaches first construct all event trends and then aggregate them

Flink is a popular open-source streaming engine that supports event pattern matching but not Kleene closure. Thus, we flatten our queries.

https://ink.apache.org/

SASE supports both Kleene closure and aggregation but does not optimize aggregation of Kleene matches.

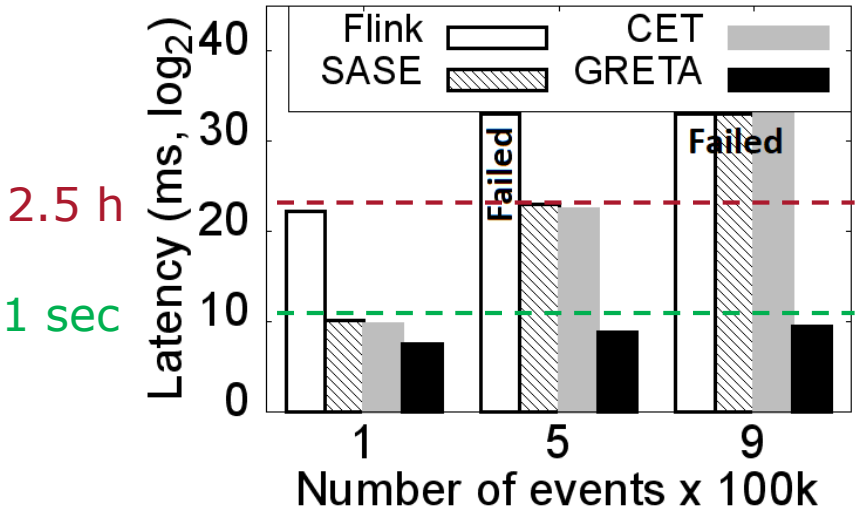
H.Zhang, Y.Diao, and N.Immerman. On complexity and optimization of expensive queries in Complex Event Processing. In SIGMOD, pages 217-228, 2014.

CET finds the middle ground between CPU time and memory usage of event trend detection. It does not support aggregation of event trends.

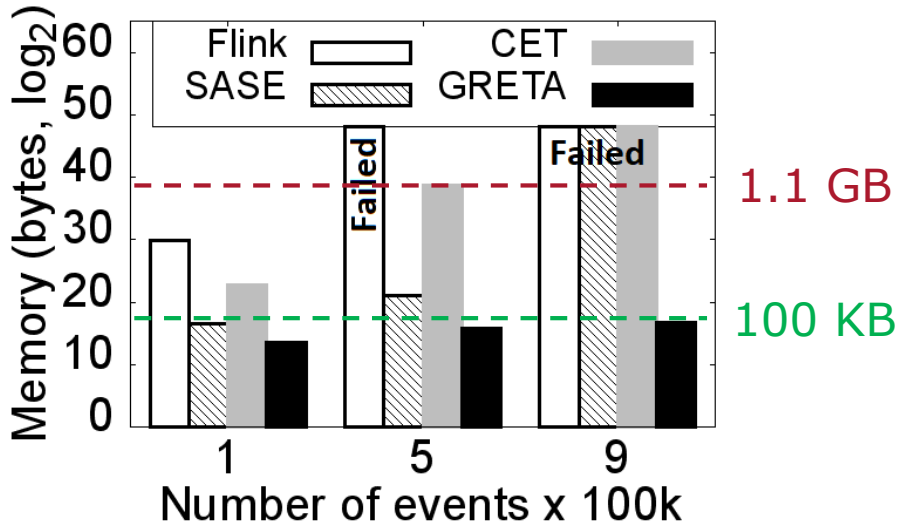
O.Poppe, C.Lei, S.Ahmed, and E.A.Rundensteiner. Complete Event Trend Detection in High-Rate Event Streams. In SIGMOD, pages 109-124, 2017.

Event Trend Aggregation

Latency



Memory



GRETA is a win-win solution that

- achieves **4 orders of magnitude speed-up** compared to all existing approaches and
- uses **50-fold less memory** than SASE

Contributions

We are the first to compute **aggregation of Kleene closure matches over event streams** with **optimal time complexity**

1. **GRETA graph** compactly encodes all event trends matched by expressive Kleene queries
2. Graph-based event trend aggregation with **quadratic time complexity**
3. **4 orders of magnitude speed-up** and **8 orders of magnitude memory reduction** compared to existing approaches

Questions?


thank you
danke
謝謝
ngiyabonga
teşekkür ederim
gracias
merci
obrigado
dziękuję
bedankt
go raibh maith agat
mochchakkeram
sukriya
kop khun krap
terima kasih
grazie
arigatō
dakujem
takk
taiku
sagolun
hvala
mauruuru
спасибо
рахмат
Баярлалаа
dank je
tapadh leat
merci
мерси
ευχαριστώ
감사합니다
ありがとう
তোমাকে ধন্যবাদ
enkosi
sobodi
dankie
kiitos
dhanyavad
nandri
nanni
bayarlalaa
gracie
köszönöm
akun
dankon
aciū
chnorakaloutioun
gratias ago
gràcies
sulpáy
djiere dieuf
tau
дякую
mamnun
chokrane
murakoze
tenki
obrigada
asante
manana
хвала
paldies
grazzi
mahalo
spas
tack
welalin
barka
mersi
kia ora
vinaka
спасибі
blagodaram
misaotra
matondo
faafetai lava
didi madloba
najis tuke
kam sah hamnida
rahmat
rahmet
tanemirt
diolch
dhanyavadagalu
shukriya
merce
xièxie
trugarez
mèsi
dèkuji