Outliers Detection using the Instantaneous Degree within IP Traffic Modelled as a Stream Graph

Audrey Wilmet¹, Matthieu Latapy¹, Robin Lamarche-Perrin²

¹Laboratoire d'informatique de Paris 6 (LIP6) - Complex Networks Team http://www.complexnetworks.fr/ ²Institut des Systèmes Complexes Paris Île de France (ISC-PIF)

RESCOM 2018, Toulouse, January 18, 2018







Introduction Context and Goals

Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

Difficulties

Our Method

Homogeneity
Detection

Identification Removal

Conclusion

Context and Goals

• IP traffic:

Flow of data across the Internet:

- → nodes: IP addresses;
- → interactions: packet exchanges;
- ightarrow temporal interactions.
- Goals:
- ① Detect outliers = irregular patterns:
- → global high activity of an IP address;
- → sharp variation in an the activity of an IP address;
- → high activity at a particular moment.
- 2 Identify them: find their exact cause in the data;
- 3 remove them precisely from the traffic.

Introduction Context and Goals

Stream Graph

Instantaneous Degree Detect. Identify.

In Practice

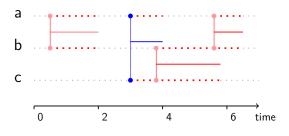
Degree Heterogeneity Difficulties

Our Method

Conclusion

IP Traffic as a Stream Graph

Stream graph: 1h of real IP Traffic (MAWI)



ex: a and c interacted from $t_1 = 3$ to $t_2 = 4$

- → M. Latapy et al., 2017; T. Viard et al., 2017.
- Other Data Modelling: Signal, graph
- → Y. Himura et al., 2013 ; H. Asai et al., 2014.

RESCOM 2018 2 / 22

Context and G Stream Graph

Our Approach

Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogene Difficulties

Our Method

Homogene

Detection Identification

Removal

Conclusion

Our Approach

Data Modelling: temporal interactions = stream graph

Outlier = entity statistically deviating from others

- 1) Entity x: (uv, t) uv (v, t) v t $\Rightarrow (v, t)$
- 2) Studied set $x \in X$:

All entities of the same types or a subset of it?

$$\psi \qquad \qquad \psi \\
t \in X = T \qquad \qquad t \in X \subset T$$

3) **Feature** \mathcal{F} on X

⇒ Instantaneous degree

Context and Goals Stream Graph

Instantaneous Degree

Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

Temporal Homogene

Detection

Identificat Removal

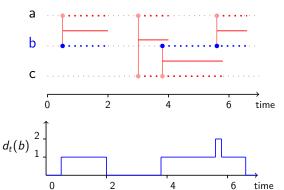
Conclusion

Instantaneous Degree of (v, t)

Number of neighbours of node *v* at time *t*:

$$d_t(v) = |\{u, (t, uv) \in E\}|$$

Example: degree profile of b



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph Our Approach

Our Approach Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

Temporal Homogeneity

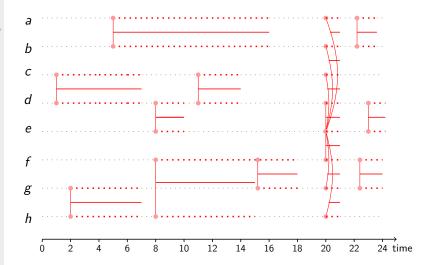
Detection

Identification &

Conclusion

1 Detection

Entity: (v, t) **Feature:** $d_t(v)$ **Studied set:** $V \times T$



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph

Our Approach
Instantaneous Degree
Detect, Identify,
Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

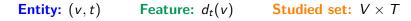
Temporal Homogene

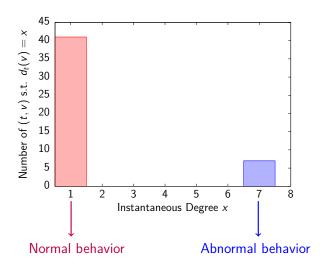
Detection

Removal

Conclusion

① Detection





Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph Our Approach

Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

Temporal Homogeneity

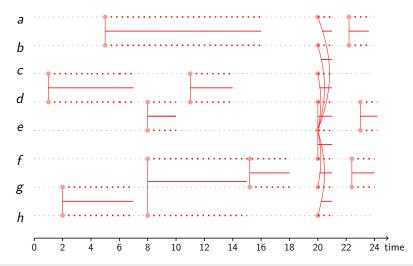
Detection

Identification &

Conclusion

① Detection

$$\{(v,t): d_t(v)=7\} \Rightarrow \text{Detected outlier}$$



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph Our Approach

Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

Temporal Homogeneity

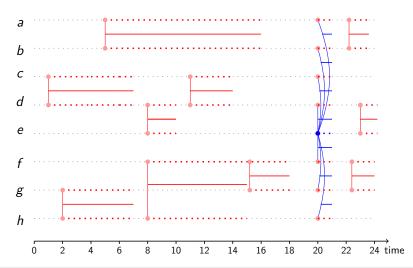
Detection

Identification & Removal

Conclusion

2 Identification

$$\{(e,t), t \in [20,21[\} \Rightarrow \text{identified outlier}\}$$



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph

Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

Temporal

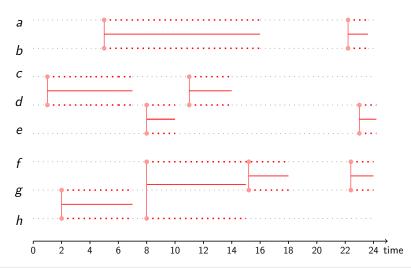
Homogeneity Detection

Identification &

Conclusion

3 Removal

$$\{(e,t),t\in[20,21[\}$$



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph

Our Approach

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

Homogeneity

Detection

Identification &

Conclusion

In Practice

Context and Goals Stream Graph

Instantaneous Degree Detect. Identify.

In Practice

Degree Heterogeneity

Difficulties

Our Method

Tomporal

Homogene Data atian

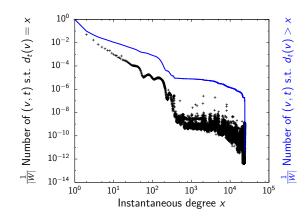
Identificat

Conclusion

Global Degree Distribution

Entity: (v, t) **Feature:** $d_t(v)$

Studied Set: Global, $W = V \times T$.



Context and Goals Stream Graph

Instantaneous Degree
Detect. Identify.

In Practice

Degree Heterogeneity
Difficulties

Our Method

Temporal Homogene

Detection

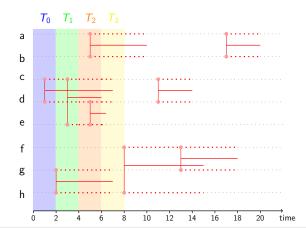
Removal

Conclusion

Local Degree Distribution (1/2)

Entity: (v, t) **Feature:** $d_t(v)$

Studied Set: Local, $W_i = T_i \times V$, $T_i = [2i, 2i + 2[...]$



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph

Instantaneous Degree
Detect, Identify,
Remove

In Practice

Degree Heterogeneity Difficulties

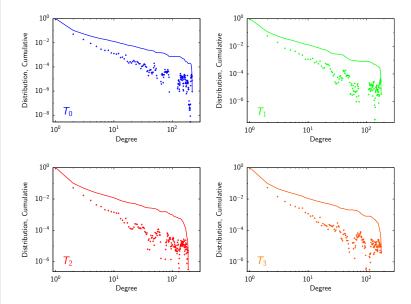
Our Method

Temporal Homogeneit Detection

Identification Removal

Conclusion

Local Heterogeneity (2/2)



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

III I Tactice

Degree Heterogeneity Difficulties

Our Method

- vietii

Homogene

Identificat

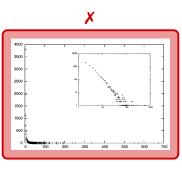
Removal

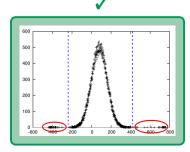
Conclusion

Difficulties

 $Outlier = Activity \ that \ deviates \ from \ the \ usual \ one$

Find an outlier \iff Find the normality





Heterogeneous

Homogeneous with outliers

Context and Goals Stream Graph

Our Approach

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

Homogeneity

Detection

Identification & Removal

Conclusion

Our Method

Context and Goals Stream Graph

Instantaneous Degree
Detect. Identify.

In Practice

Degree Heterogeneity Difficulties

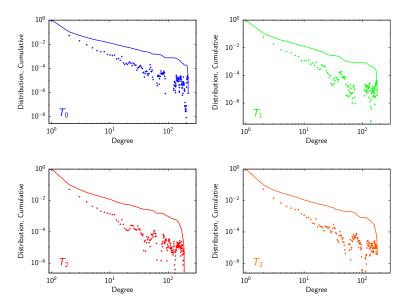
Our Method

Temporal Homogeneity

Detection Identification

Conclusion

Temporal Homogeneity



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph Our Approach

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

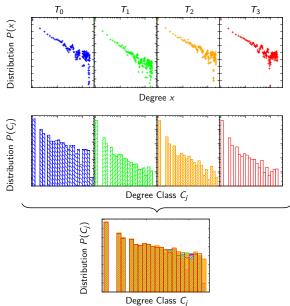
Tempor

Homogeneity Detection

Identification &

Conclusion

Comparison of Local Distributions



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph Our Approach

Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

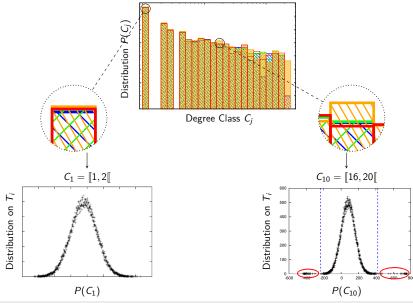
Temporal Homogeneity

Detection

Identification & Removal

Conclusion

Comparison of Local Distributions



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

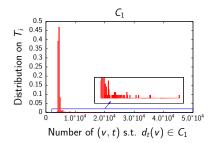
Temporal

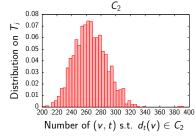
Detection

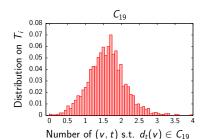
Identificatio

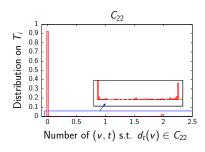
Conclusion

Results: Homogeneous Distributions









Context and Goals Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

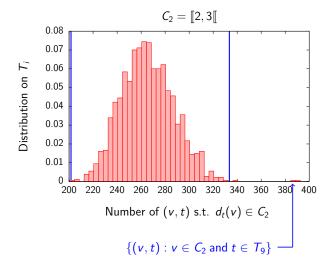
Temporal

Detection

Identificat

Conclusion

Results: Homogeneous Distributions



Outlier = temporal T_i + structural C_i information.

Context and Goals Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

Temporal

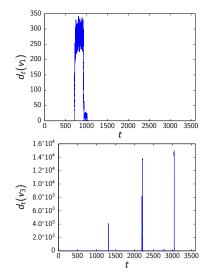
Detection

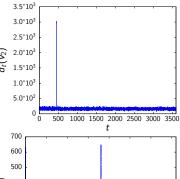
Identification & Removal

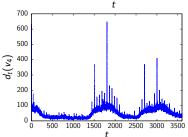
Conclusion

Identification and Removal - Degree Profiles

 \implies Suspicious couples (v, T_i) .







Context and Goals Stream Graph Our Approach

Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

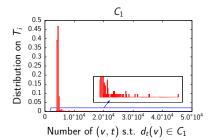
Temporal Homogene

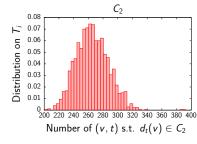
Detection

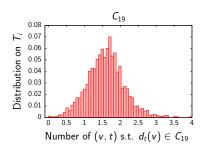
Identification & Removal

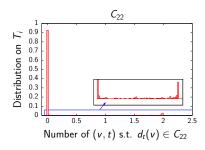
Conclusion

Distributions after removals









Context and Goals Stream Graph

Our Approach
Instantaneous Degree
Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

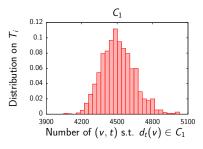
Temporal

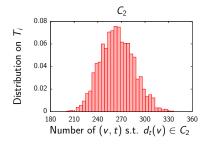
Detection

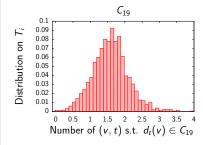
Identification & Removal

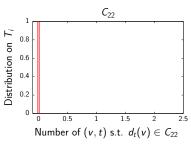
Conclusion

Distributions after removals









Introduction Context and Goals

Stream Graph Our Approach

Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

.

Our Method

Homogene

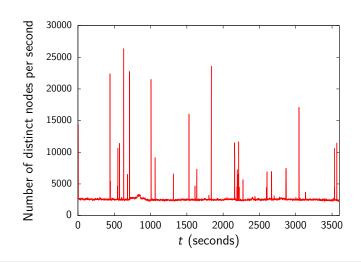
Detection
Identification &

Removal

Conclusion

Removal of identified outliers (v, T_i)

 \Longrightarrow Consequence on the number of distinct nodes per second.



Context and Goals Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

. .. .

Our Method

Tomoroul

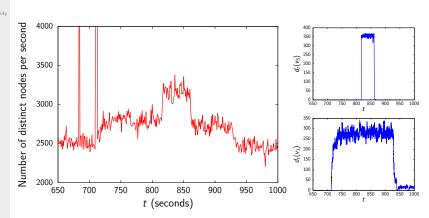
Homogeneity

Identification &

Conclusion

Removal of identified outliers (v, T_i)

 \Longrightarrow Consequence on the number of distinct nodes per second.



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

0 14 .1

Our Method

- ...

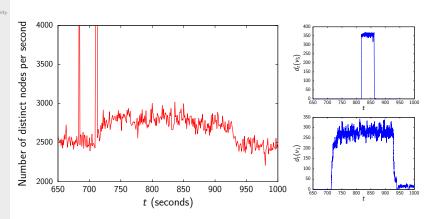
Homogeneit

Detection
Identification &

Removal Conclusion

Removal of identified outliers (v, T_i)

 \Longrightarrow Consequence on the number of distinct nodes per second.



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Context and Goals Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

. .. .

Our Method

Our wicting

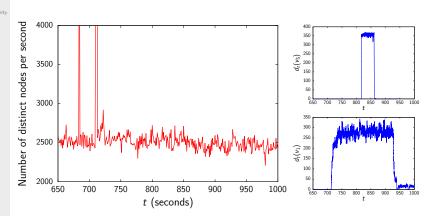
Homogeneity

Detection Identification & Removal

Conclusion

Removal of identified outliers (v, T_i)

 \Longrightarrow Consequence on the number of distinct nodes per second.



Audrey Wilmet, Matthieu Latapy, Robin Lamarche-Perrin

Introduction Context and Goals

Stream Graph

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

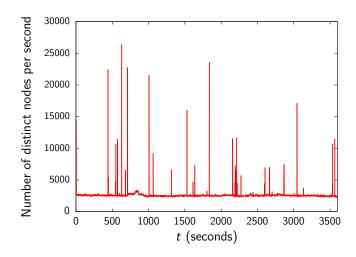
Identification &

Conclusion

Removal

Removal of identified outliers (v, T_i)

Consequence on the number of distinct nodes per second.



Context and Goals Stream Graph Our Approach

Instantaneous Degree Detect, Identify, Remove

In Practice

Degree Heterogeneity Difficulties

Our Method

Temporal

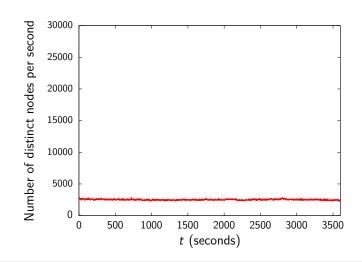
Homogenei

Identification & Removal

Conclusion

Removal of identified outliers (v, T_i)

 \Longrightarrow Consequence on the number of distinct nodes per second.



Context and Goals Stream Graph

Our Approach

Instantaneous Degree Detect, Identify,

In Practice

Degree Heterogeneity Difficulties

Our Method

Homogeneity

Detection

Identification & Removal

Conclusion

Conclusion

RESCOM 2018 20 / 22

Stream Graph
Our Approach
Instantaneous Degree
Detect. Identify.

In Practice

Degree Heterogeneity Difficulties

Our Method

Homogeneity
Detection
Identification &

Conclusion

Conclusion

- Contributions:
- Modelling of IP traffic as a stream graph
- Design of a method to detect outliers in heterogeneous distributions
- \rightarrow IP with anomalous degree profile, network scans.
- Iterative removal of identified outliers
- \rightarrow Return to normal traffic (w.r.t $d_t(v)$).
- ⇒ Method applicable over temporal interactions in general.
- Several possible improvements:
- Identification of couples (v, t) instead of couples (v, T_i) ,
- more complex feature: clustering coefficient,
- exploring other assumptions:
- → T. Schieber et al., Q. Zhang et al., 2017.

Context and Goals Stream Graph

Instantaneous Degree Detect, Identify.

In Practice

Degree Heterogeneity Difficulties

Our Method

Temporal Homogeneity

Detection Identification &

Conclusion

Thanks for your attention !

RESCOM 2018 22 / 22