

# Temporal Graph Algebra

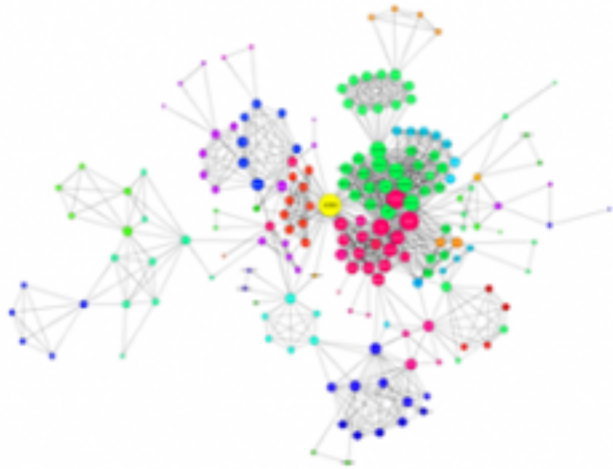
---

VERA ZAYCHIK MOFFITT

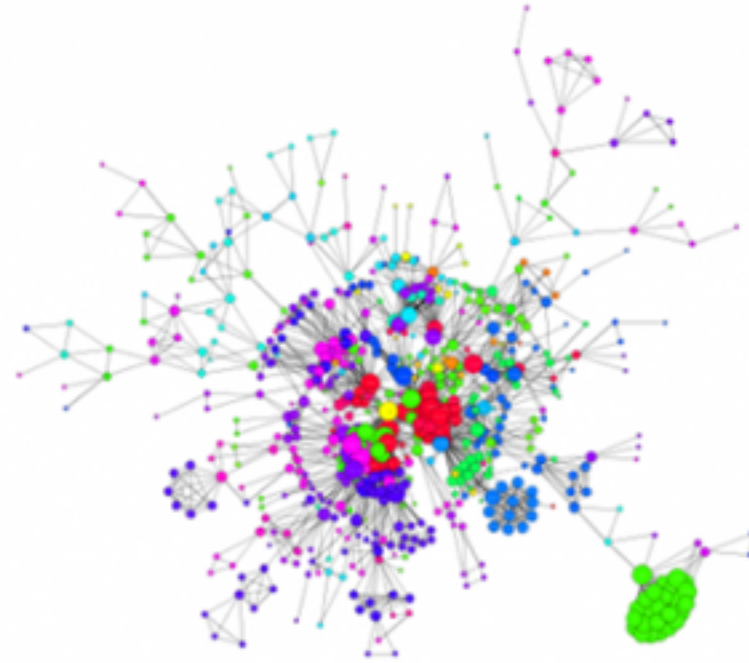
JOINT WORK WITH JULIA STOYANOVICH

SEPTEMBER 1, 2017

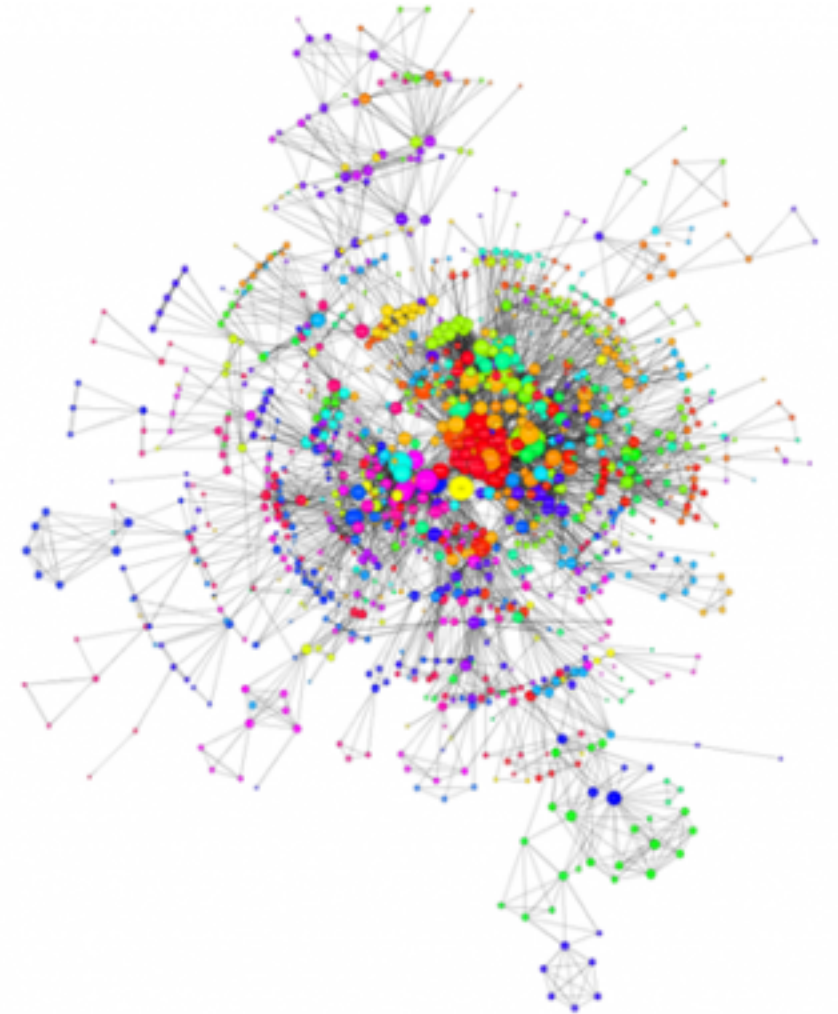
# Graph Evolution



**2007-2008**



**2009-2010**



**2011-2012**

# Interesting and Important Questions

---

- What is the likelihood of an individual to join a community?
- Which roads exhibit abrupt congestion and at what time?
- Which websites have the highest increase in popularity/rank over the past year?
- What is the rate of densification of the graph?
- Have any changes in network connectivity been observed?
- At what time scale can interesting trends be observed?

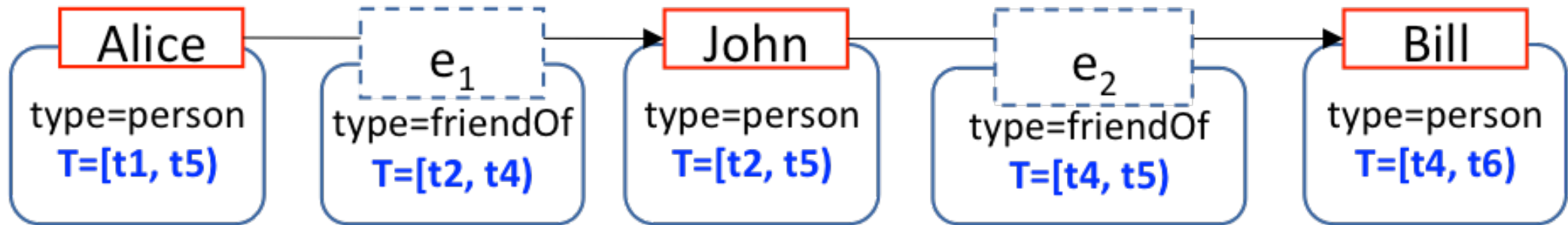
# Goal

---

**Principled and systematic** support for querying  
and analytics of evolving graphs

# Existing Models – Time as Data

Approach – Add time property



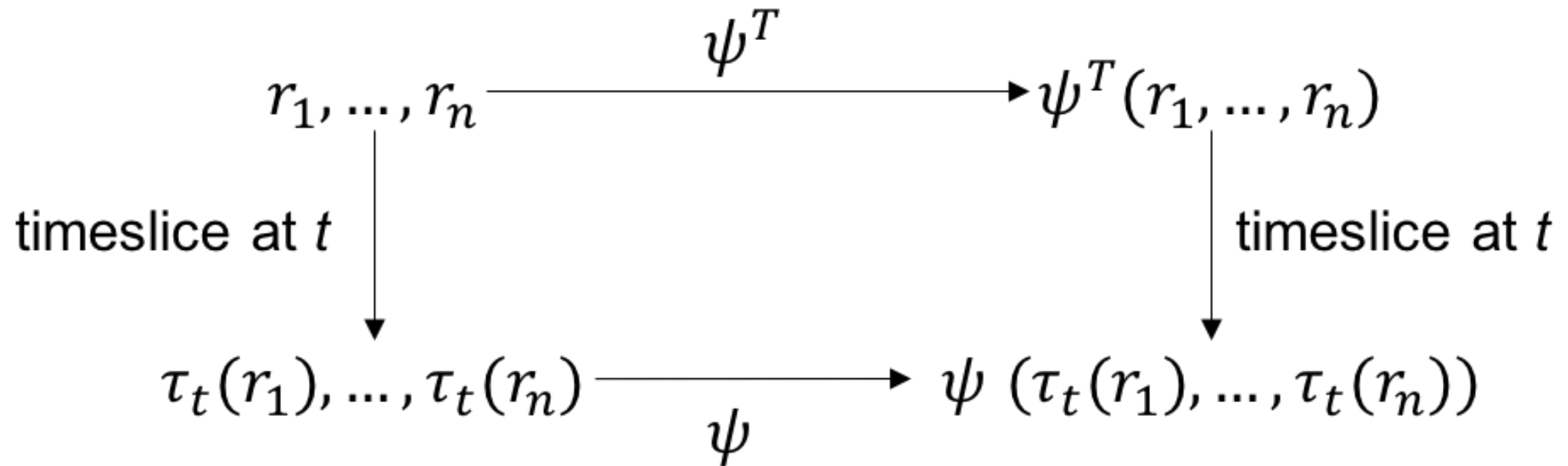
- Need a new node for each change of property or period of validity

Are Alice and Bill connected?

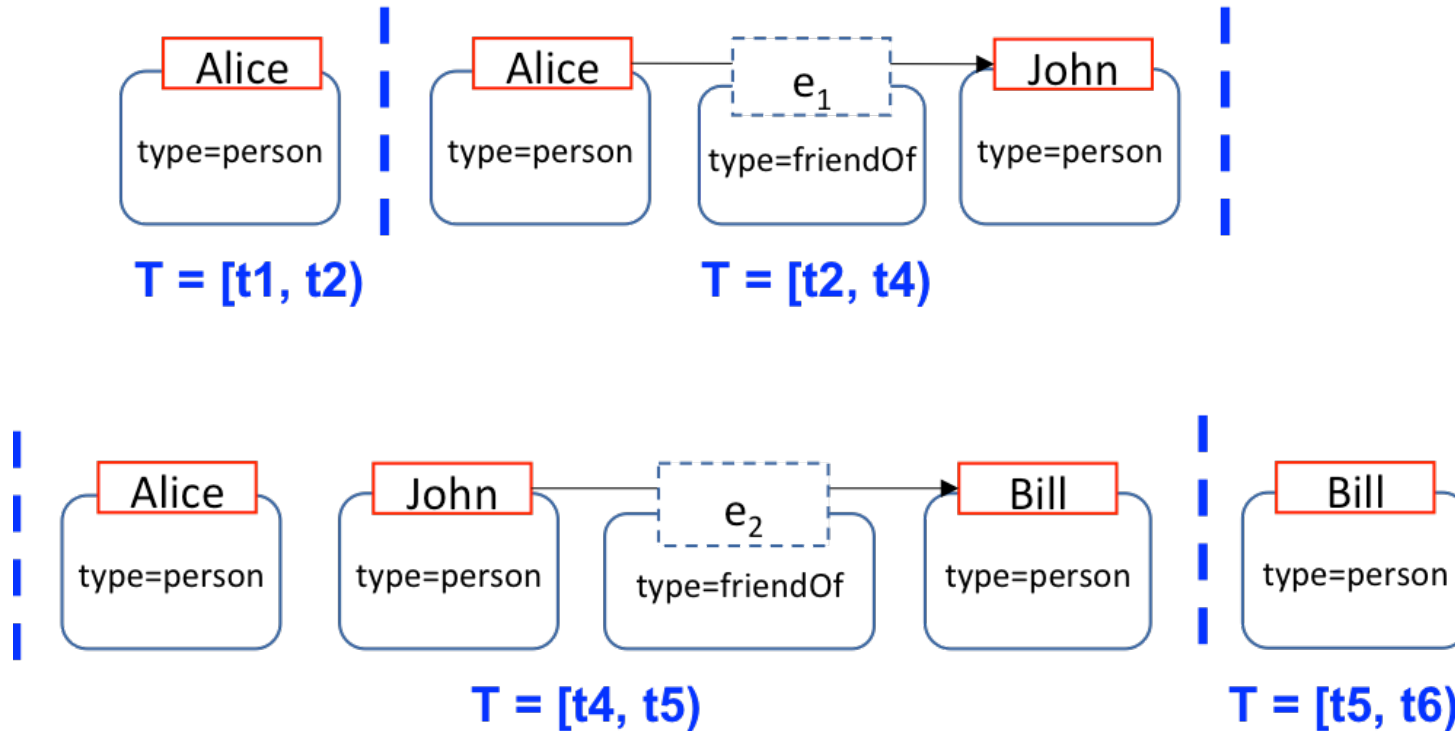
**! Time needs special treatment**

# Snapshot Reducibility

---



# Existing Models – Snapshot Sequence



Which pairs of people are connected by a journey?

**! No explicit references to time**

# Contributions

---

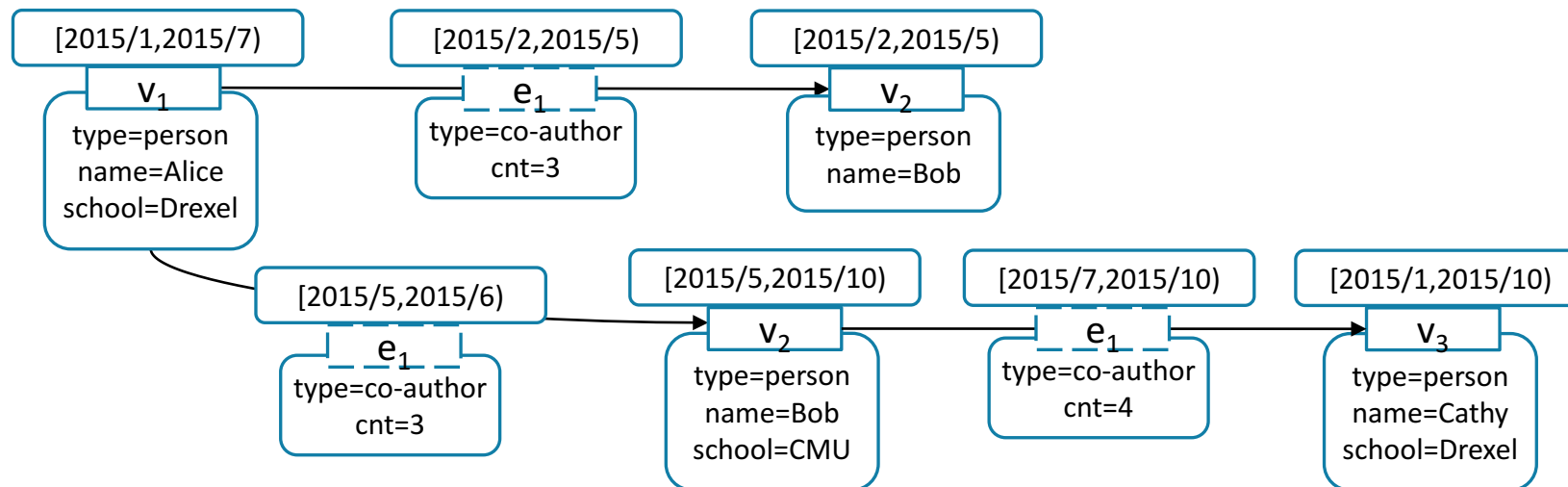
- Conceptual representation of an evolving graph
  - Captures evolution of both topology and properties
- Temporal Graph Algebra (TGA)
  - Concisely express wide range of common analysis tasks



# Temporal Graph Model

Definition 3.1.1 (TGraph). A TGraph  $\mathcal{G}$  is a 7-tuple  $(V, E, \Pi, \rho, \xi, \lambda_v, \lambda_e)$ , where:

- $V$  – set of nodes,
- $E$  – set of edges,
- $P$  – set of available properties,
- $\rho : E \rightarrow (V \times V)$  total function,
- $\xi^T : (V \cup E) \times T \rightarrow B$  total function,
- $\lambda^T : (V \cup E) \times P \times T \rightarrow Val$  partial function



# TGA Operators

---

- Provide temporal versions of common graph operations:
  - subgraph
  - aggregation
  - vertex- and edge-map
  - union, intersection, difference

# Node Creation

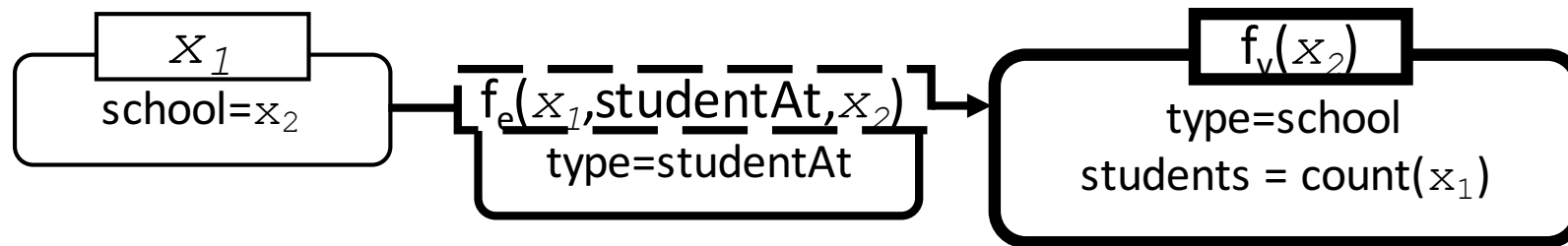
---

- attribute-based node creation
  - add new nodes representing a matching input pattern
- window-based node creation
  - Change temporal resolution of  $\mathcal{G}$

# Attribute-based node creation

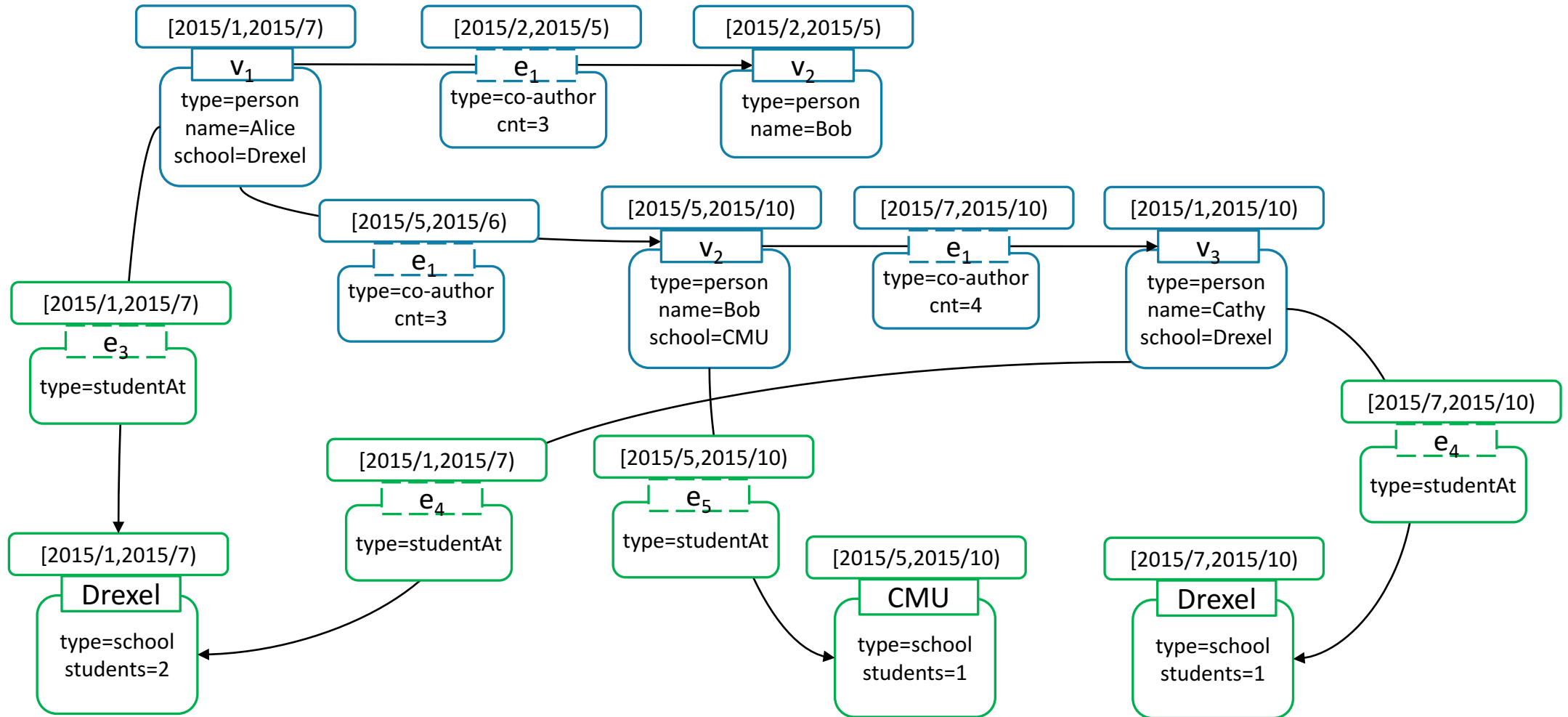
---

- Add new nodes to represent matching pattern



adds nodes Drexel and CMU and edges to them

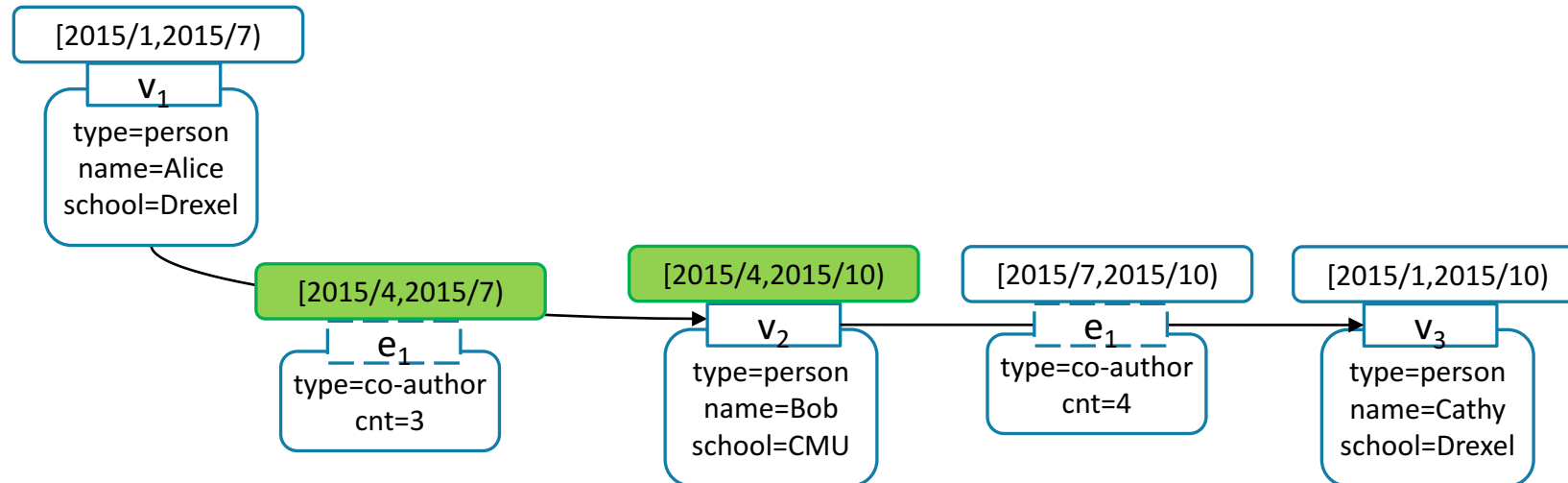
# Attribute-based node creation



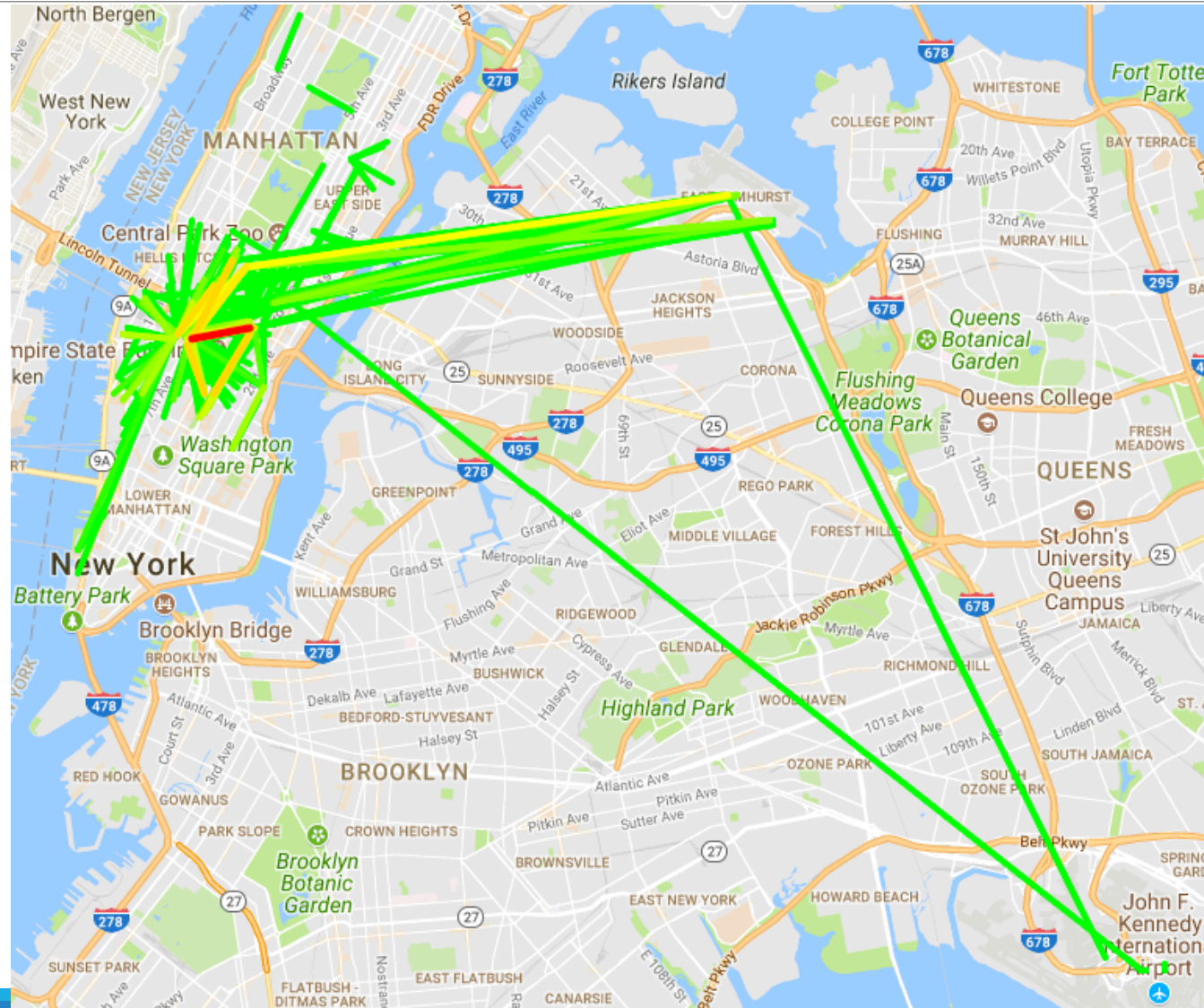
# Window-based node creation

$\text{node}_w^T(r_v = \text{always}, r_e = \text{exists}, f_{v_1} = \text{first}(\text{name}), f_{v_2} = \text{first}(\text{school}), \mathcal{G})$

3 months

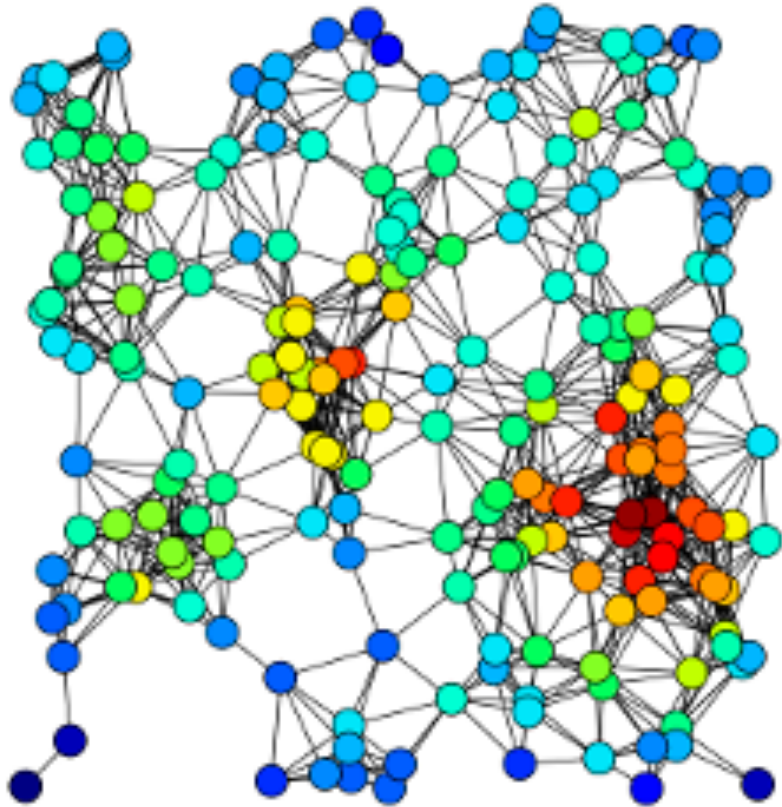


# Example: NYC Cabs



# Node Influence over Time

---



Are there high influence nodes and is that behavior persistent over time?



# Node Influence, with TGA

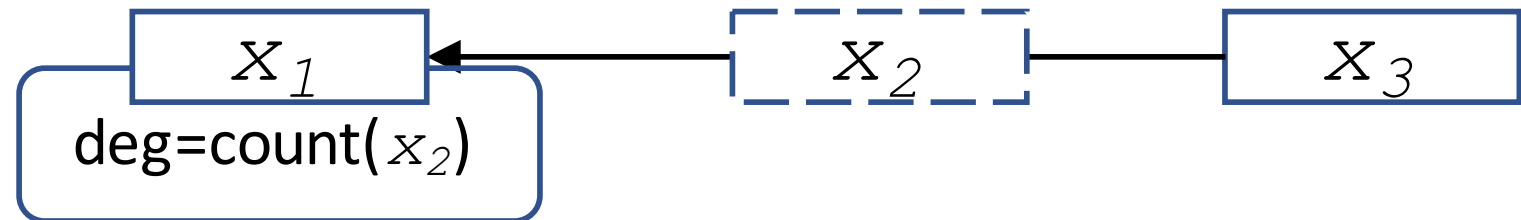
---

1. Select a subset of the data representing the 5 years of interest, using trim:

$$\mathcal{G}_1 = \text{trim}_{[2010,2015]}^T(\text{wikitalk})$$

2. Compute in-degree (prominence) of each node during each time point using aggregation and pattern p1

$$\mathcal{G}_2 = \text{agg}_{p_1}^T(\mathcal{G}_1)$$



# Node Influence, with TGA

---

3. Aggregate degree information per node across the timespan of  $\mathcal{G}_2$  using the window-based node creation operator:

$$\mathcal{G}_3 = \text{node}_w^T(w = \text{lifetime}, f_v = \{\text{map}(\text{deg})\}, \mathcal{G}_2)$$

4. Transform the attributes of each node using the vertex-map operator:

$$\mathcal{G}_4 = \text{map}_v^T(f_v = \text{stdev}(\text{deg})/\text{mean}(\text{deg}) * 100, \mathcal{G}_3)$$

# Summary

---

- TGraph model represents evolution of graph topology and properties
- TGA provides a concise set of operations over TGraphs
  - Precise semantics
  - More expressive than current state of the art
  - Desirable temporal properties

Thank You!

Questions?

