Graph Macro Dynamics with Self-Attention for Event Detection

Mert Kosan[1], Arlei Silva[2], Sourav Medya[3], Brian Uzzi[4], Ambuj Singh[1]

[1]University of California, Santa Barbara
[2]Rice University
[3]University of Illinois Chicago
[4]Northwestern University

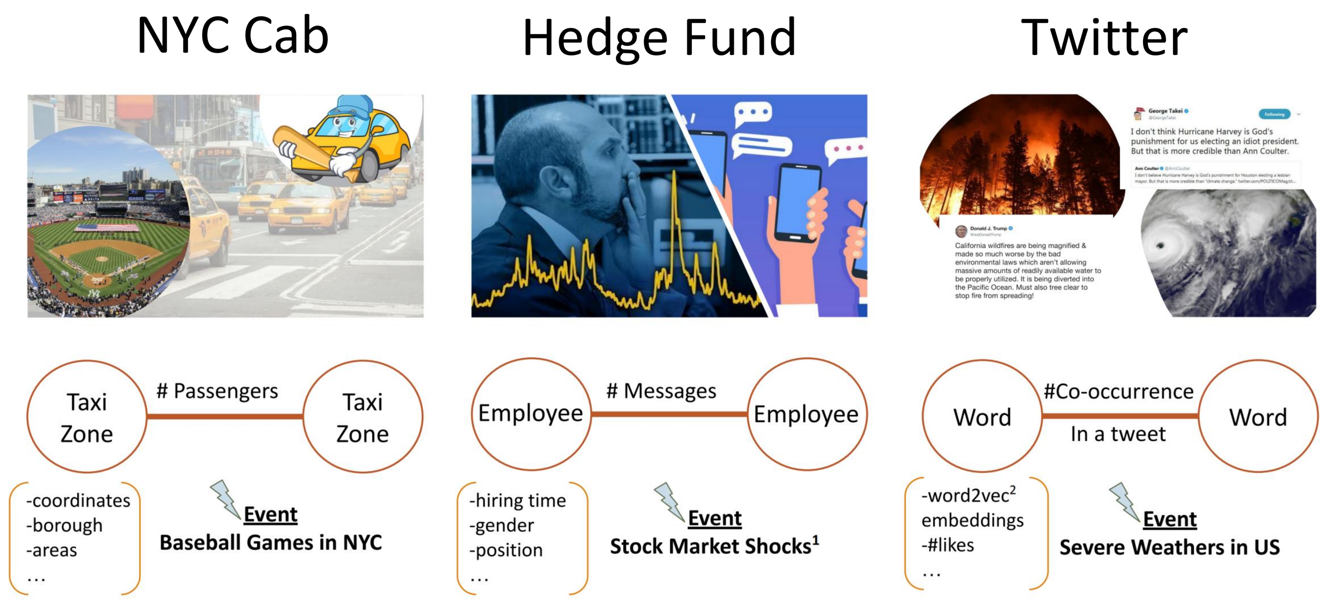
Abstract

We propose DyGED, which learns correlations between the graph macro dynamics (i.e., a sequence of graph-level representations) and labeled events. Our approach combines structural and temporal self-attention to account for application-specific node and time importances effectively.

Contributions

- First study comparing micro and macro architectures for event detection.
- Structural and temporal self-attention enables the effective learning of node and time dependent weights.
- Comparison against several baselines using three datasets, outperforming them by up to 8.5% while being scalable.

External Events



	NYC Cab	Hedge Fund	TW	TW-Large
#Nodes (avg)	263	330	300	1000
#Edges (avg)	3717	557	1142	10312
#static features	6	5	300	300
#Snapshots	4464	690	2557	2557
Snap. Period	hour	day	day	day
#Events	162	55	287	287

Architecture

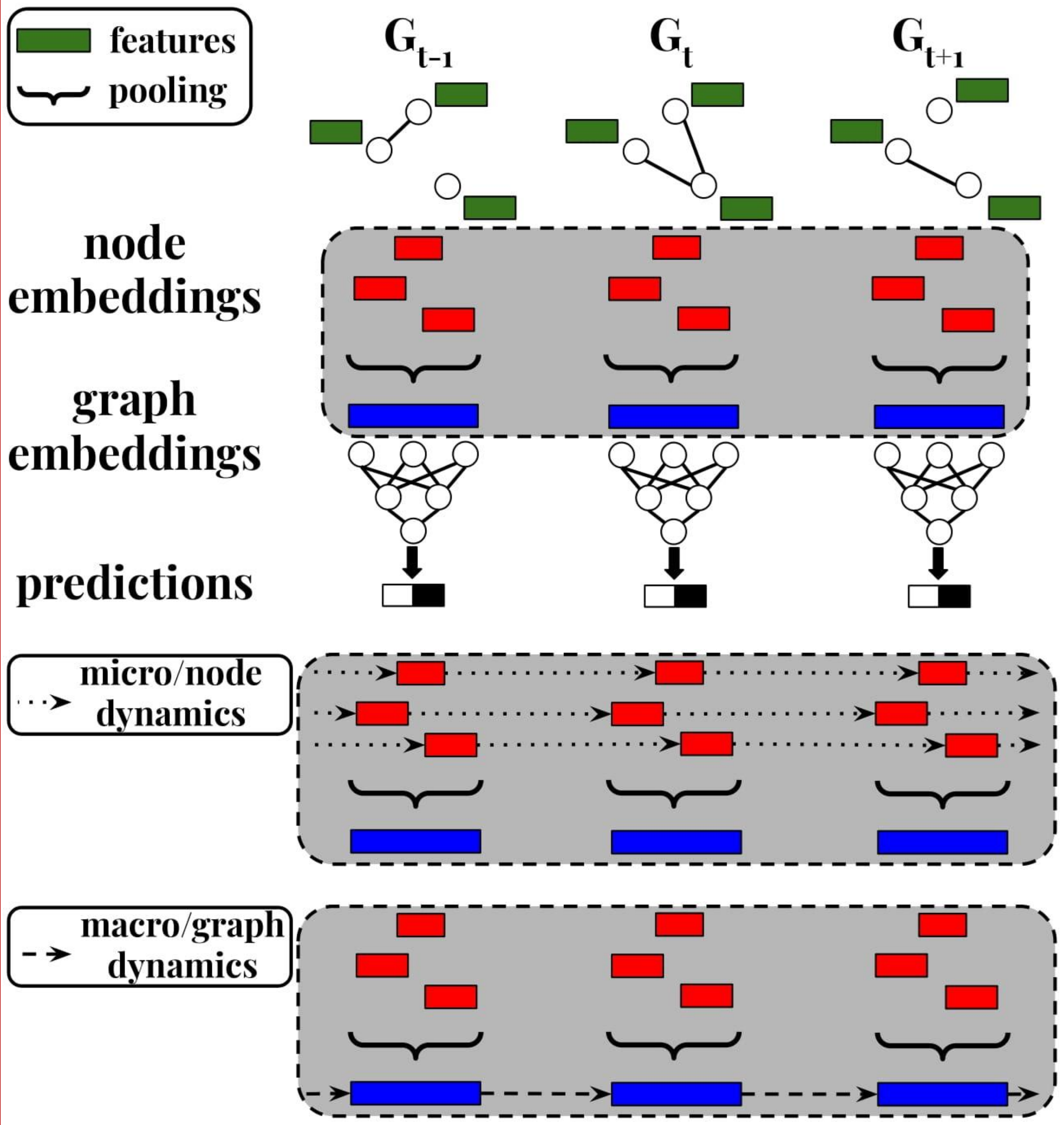


Figure: (a) Event detection on dynamic graphs based on a deep learning architecture. (b) At the micro scale, the dynamics is captured at the node-level. (c) At the macro scale, the dynamics is captured at the graph level.

DyGED Algorithm

Require: Sequence of snapshots $G_{t-k:t}$, previous dynamic state z'_{t-k-1}
Ensure: Event probability

for $\tau \in \{t-k, \dots, t\}$ **do**
  $Z_\tau \leftarrow GCN(G_\tau, X_\tau)$
  $z_\tau \leftarrow$ v-Att(Z_τ)
  $z'_\tau \leftarrow LSTM(z'_{\tau-1}, z_\tau)$
end for
$z''_t \leftarrow$ t-Att$([z'_{t-k}; \dots; z'_t])$
return $MLP(z''_t)$

Self-attention mechanism [3]

$$\text{v-Att}(Z_t) = \mathbf{softmax}(\mathbf{w}.\tanh(\Phi Z_t^T))Z_t$$

$$\text{t-Att}(Z'_t) = \mathbf{softmax}(\mathbf{w}'.\tanh(\Phi' Z_t'^T))Z'_t$$

Experiments

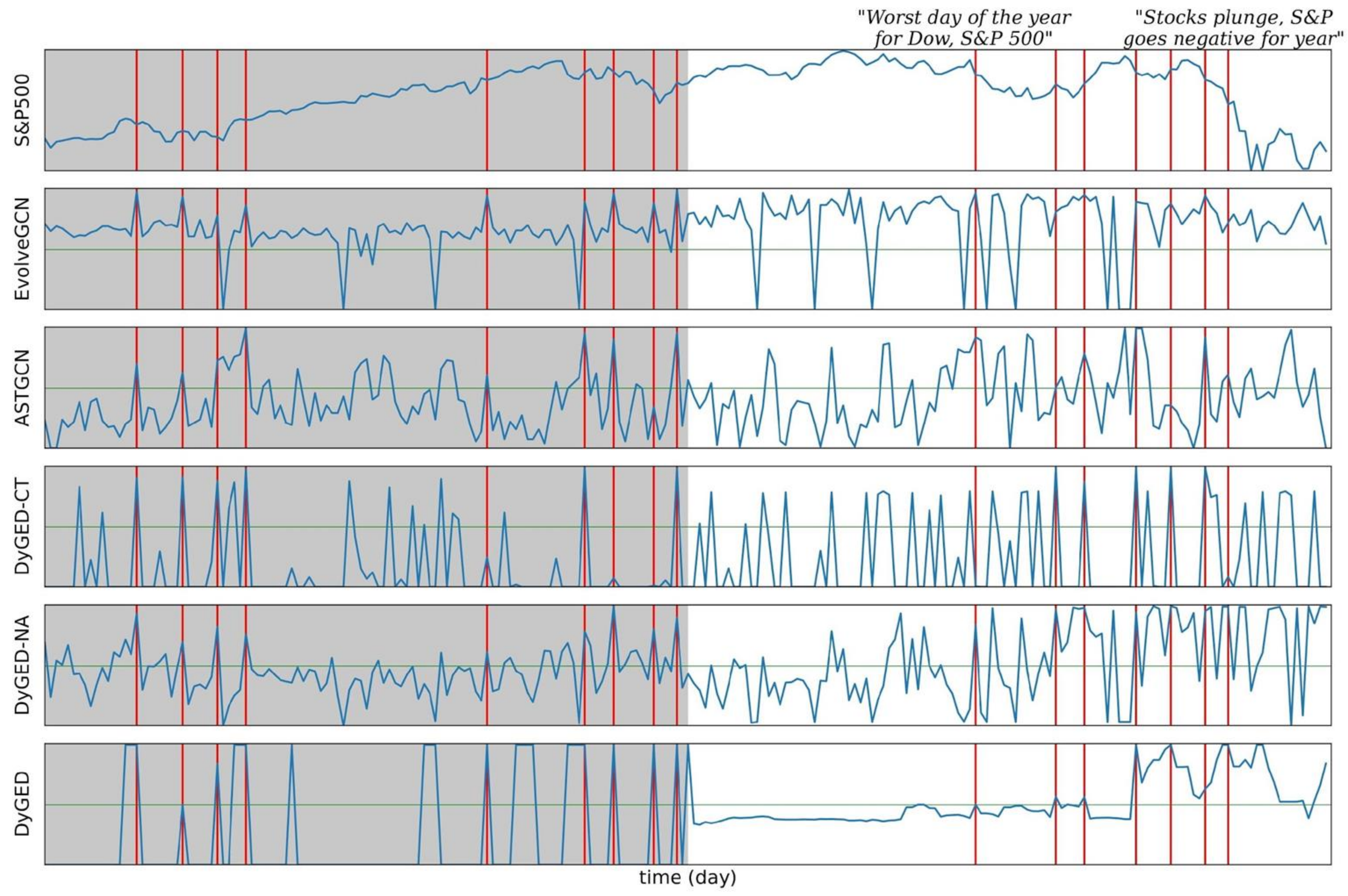
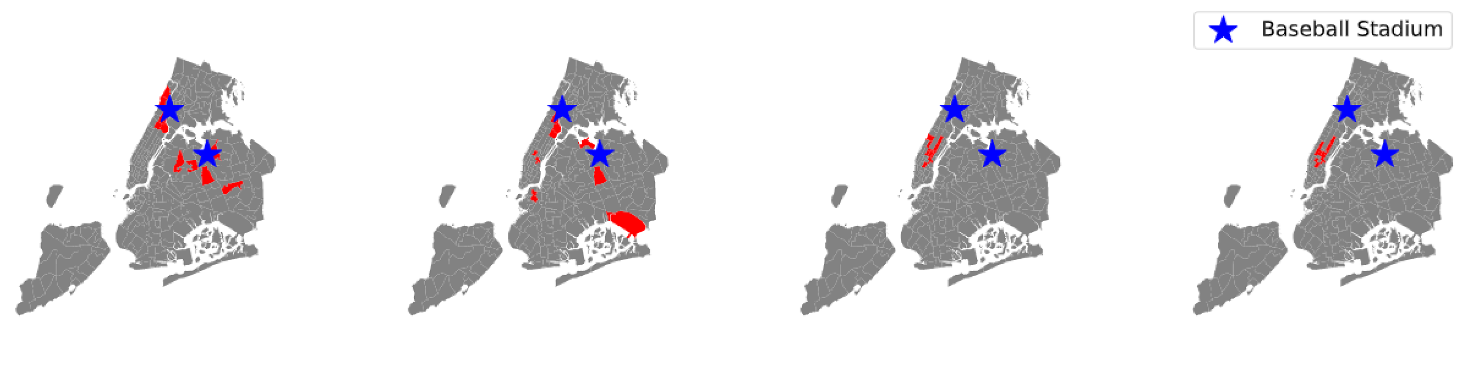
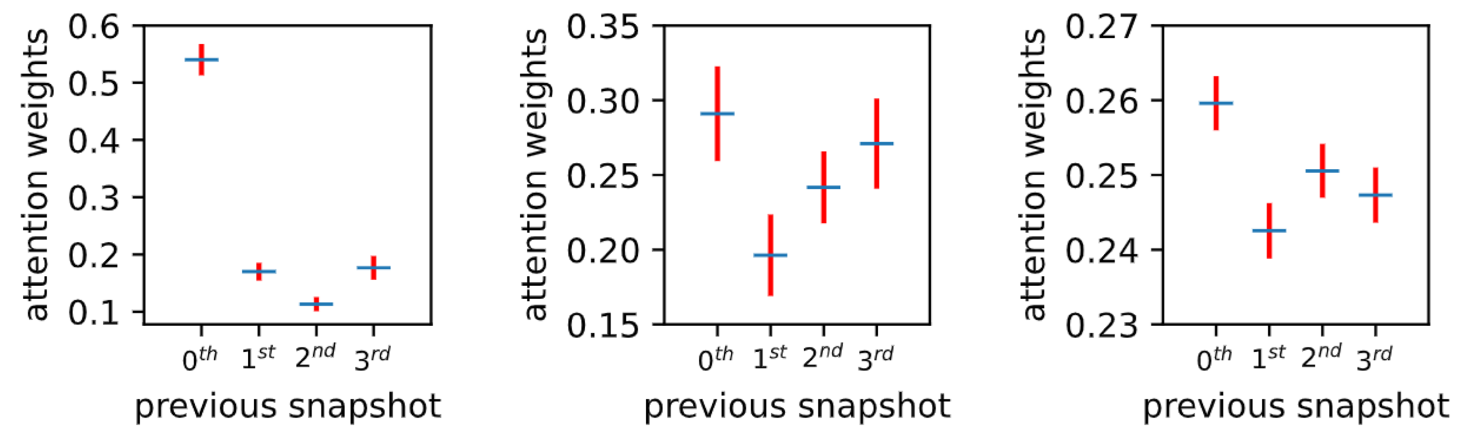


Figure: Scaled prediction by event detection methods for Hedge Fund. Events are marked as vertical (red) lines. DyGED identifies several of the events while keeping the false positive rate lower.



(a) DyGED Attention (b) Betweenness (c) Clustering (d) Degree

Figure: Top 10 taxi zones (in red) based on various metrics for NYC Cab dataset. DyGED Attention mechanism finds closer taxi zones to the stadiums compared to some node statistics.



(a) NYC Cab (b) Hedge Fund (c) Twitter Weather

Figure: Illustration of the normalized attention weights of current (0th) and previous three snapshots for all datasets. Results show the history plays a role in predicting the event.

References:

[1] Daniel M. Romero, Brian Uzzi, Jon Kleinberg. Social Networks Under Stress. Web Conference 2016
[2] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality. NeurIPS 2013
[3] Li, J.; Rong, Y.; Cheng, H.; Meng, H.; Huang, W.; and Huang, J. 2019. Semi-supervised graph classification: A hierarchical graph perspective. In The Web Conference, 972–982.

Code: https://github.com/mertkosan/DyGED