Прогнозирование движения объектов с применением графа трехмерной сцены

Трунова А.М.
ttrunovaa@gmail.com

Центр когнитивного моделирования МФТИ

2025 – Научно-исследовательская работа





Мотивация

Современные **автономные системы**, включая беспилотные автомобили, роботов-манипуляторов и промышленные автоматизированные платформы, требуют не только точного восприятия окружающего мира, но и способности предсказывать его изменения. Умение моделировать будущее состояние сцены позволяет таким системам **принимать упреждающие решения**, снижая вероятность аварийных ситуаций и повышая эффективность их работы, **эффективно планировать траектории движения**, учитывая возможные изменения сцены в будущем.

Цель исследования

Разработать эффективный метод предсказания будущего состояния сцены на основе графового представления объектов и их взаимодействий.

Постановка задачи

Пусть имеется последовательность наблюдений сцены во времени, представимая как $G_0, G_1, \ldots, G_t$, где:

- $G_t = (V_t, E_t)$ — граф сцены в момент времени t;
- V_t — множество объектов в момент времени t;
- E_t — множество связей между объектами, присутствующих в момент времени t.

Необходимо спрогнозировать граф сцены в момент $t + \delta$. То есть построить отображение $F : (G_0, \ldots, G_t) \to \hat{G}_{t+\delta}$, которое приближает истинную эволюцию $G_{t+\delta}$.

Постановка задачи

Обновление состояния памяти осуществляется следующим образом:

$$h_t = f_h(h_{t-1}, G_t, a_t),$$

где:

- h_t — скрытое состояние памяти в момент времени t;
- h_{t-1} — скрытое состояние памяти на предыдущем шаге;
- G_t — текущий граф;
- a_t — действие (активность) агента в момент времени t;
- f_h — функция обновления состояния памяти.

Постановка задачи

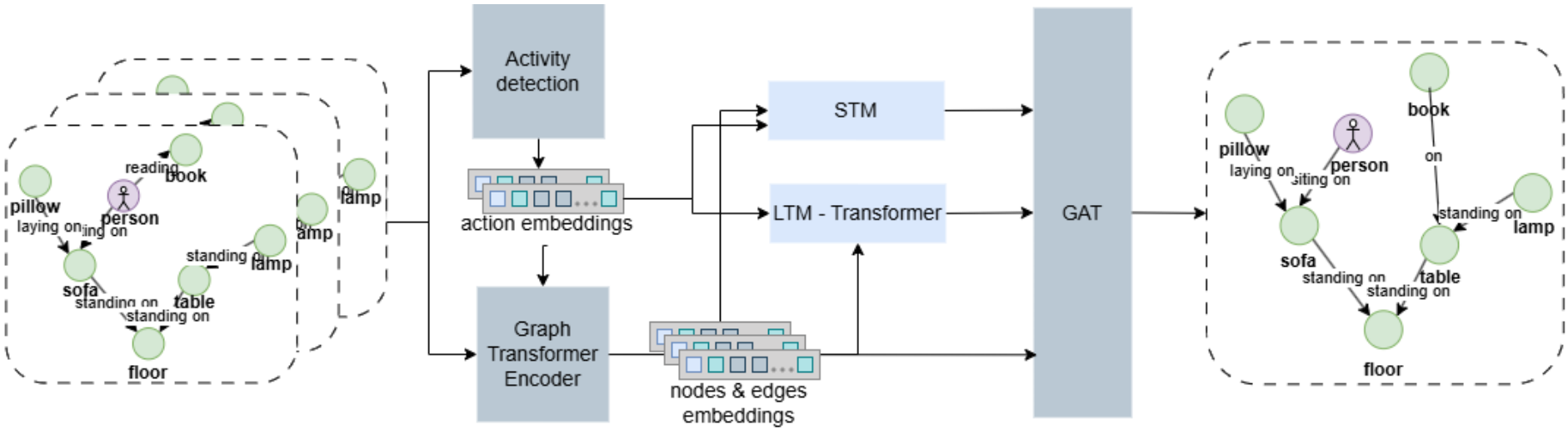
Оптимизация метода за счёт минимизации функции потерь:

$$\mathcal{L} = \mathbb{E}_t \Big[L_v(V_{t+1}, V^*_{t+1}) + L_e(E_{t+1}, E^*_{t+1}) \Big]$$

где:

- (V^*_{t+1}, E^*_{t+1}) — истинные состояния графа.
- L_v и L_e — функции потерь для узлов и рёбер.
- $\mathbb{E}_t$ — математическое ожидание по времени.

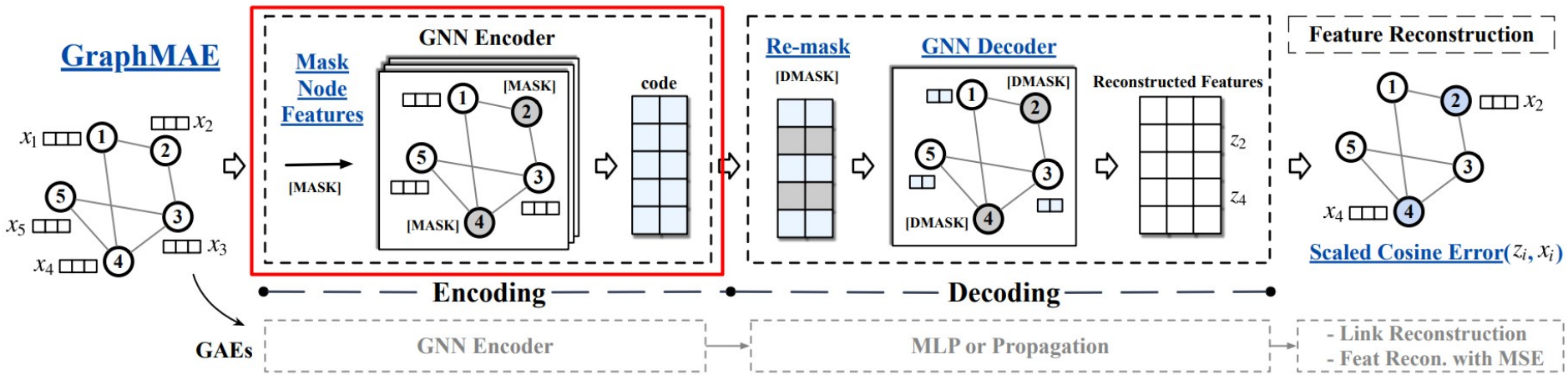
Общая схема предлагаемого подхода



Задачи

- Модификация и обучение GraphMAE;
- Выбор и модификация модели для распознования действий человека для реализации блока Activity Detection;
- Разработка и внедрение в метод механизма краткосрочной памяти;
- Разработка и внедрение в метод механизма долгосрочной памяти;
- Реализация GAT для предсказания будущего графа;
- Определение метрик для оценки модели.

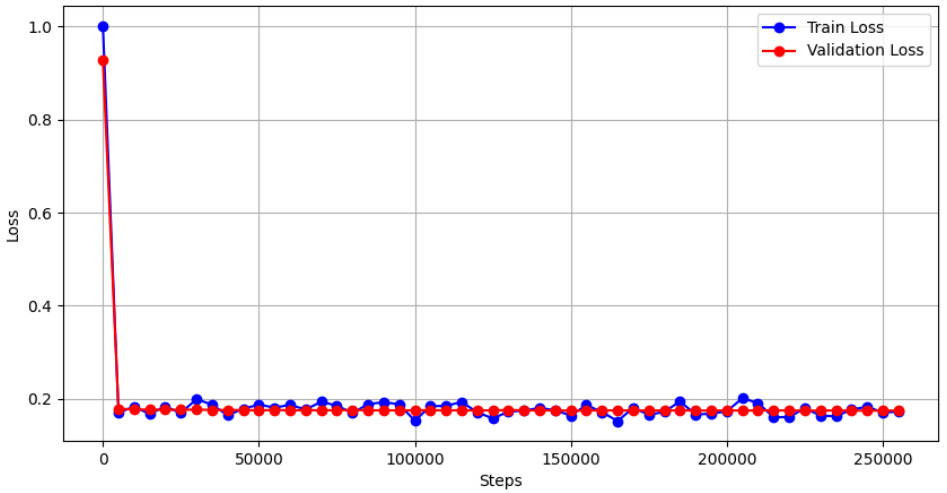
Первые эксперименты



Первые эксперименты

	GraphMAE (MUTAG)	GraphMAE (STAR Benchmark)	GraphMAE (STAR Benchmark + GraphTransformer)
Soft Contrastive Embedding Loss	0.51	0.26	**0.17**

Таблица: Сравнение значений функции потерь при использовании разных энкодеров в GraphMAE.



Спасибо за внимание!

cogmodel.mipt.ru

ttrunovaa@gmail.com