

Kỹ thuật Đặc tả: Gaussian Evolution PairRE (GE-PairRE)

Mục tiêu: Xây dựng mô hình TKGE liên tục vượt qua TNTComplex bằng cách thay thế hàm Fourier (Sin/Cos) toàn cục bằng **Gaussian Radial Basis Functions (RBF)** cục bộ. Cơ chế này đảm bảo: (1) Không gây nhiễu cho quan hệ tĩnh, (2) Bắt chính xác các xung sự kiện động, (3) Cho phép thực thể tiến hóa liên tục.

1. Triết lý Thiết kế (Core Philosophy)

- Locality (Tính cục bộ):** Thời gian chỉ tác động khi có sự kiện. Ngoài vùng ảnh hưởng, tác động thời gian tự động triệt tiêu về 0.
- Additive Evolution (Tiến hóa cộng dồn):** Thay vì dùng phép nhân (Gating) gây nhiễu, sử dụng phép cộng (Residual Offset) để thực thể "di chuyển" trong không gian nhúng theo thời gian.
- Relation-Conditioned Pulses:** Mỗi quan hệ tự học các "điểm nóng" thời gian (tâm Gaussian) của riêng mình.

2. Ký hiệu & Không gian (Notation)

- $h, t \in \mathbb{R}^d$: Vector thực thể tĩnh (Base embeddings).
- $r^H, r^T \in \mathbb{R}^d$: Vector quan hệ PairRE (Projections).
- $\tau \in [0, 1]$: Thời gian liên tục đã chuẩn hóa (BẮT BUỘC dùng miền $[0, 1]$).
- K : Số lượng nhân Gaussian (Pulses) cho mỗi chiều (ví dụ: $K = 16, 32$).

3. Gaussian Temporal Encoder (GTE)

Mục tiêu là tạo ra một vector biến thiên thời gian $\Delta e_r(\tau)$ đại diện cho sự thay đổi của thực thể dưới tác động của quan hệ r tại thời điểm τ .

3.1 Cấu trúc Nhân Gaussian (Kernel)

Mỗi quan hệ r sở hữu một bộ tham số:

- $\mu_{r,k}$: Tâm của xung thời gian (Time centers), $\in [0, 1]$.
- $\sigma_{r,k}$: Độ rộng của xung (Width/Spread).
- $A_{r,k}$: Biên độ ảnh hưởng (Amplitudes), $\in \mathbb{R}^d$.

Công thức xung thứ k :

$$G_{r,k}(\tau) = \exp \left(-\frac{(\tau - \mu_{r,k})^2}{2\sigma_{r,k}^2 + \epsilon} \right)$$

3.2 Tổng hợp Tiến hóa (Evolution Vector)

$\Delta e_r(\tau)$ là tổng chập của các xung:

$$\Delta e_r(\tau) = \sum_{k=1}^K A_{r,k} \cdot G_{r,k}(\tau) \in \mathbb{R}^d$$

Tính chất quan trọng:

- Nếu τ xa các tâm μ , $G \rightarrow 0$, dẫn đến $\Delta e_r(\tau) \rightarrow 0$.
- Điều này bảo vệ tuyệt đối các quan hệ tĩnh (Static preservation).

4. Scoring Function (Additive PairRE)

Thay vì gặt thời gian vào kết quả cuối, ta cho thực thể **biến thiên** trước khi tương tác quan hệ:

1. Evolved Entities:

$$h_{evolved} = h + \Delta e_r(\tau)$$

$$t_{evolved} = t + \Delta e_{r_inv}(\tau)$$

(Ghi chú: r_inv là quan hệ nghịch đảo để đảm bảo tính đối xứng học tập)

2. Interaction (PairRE Core):

$$x = h_{evolved} \circ r^H - t_{evolved} \circ r^T$$

3. Final Score:

$$\phi(h, r, t, \tau) = -\|x\|_1$$

5. Chiến lược Huấn luyện & Loss Functions

5.1 Primary Loss (Cross-Entropy)

Sử dụng cơ chế 1-vs-All trên toàn bộ thực thể để tối ưu hóa khả năng xếp hạng (MRR).

5.2 Hard Temporal Discrimination Loss (L_{time}) - BẮT BUỘC

Đề ép các nhân Gaussian hội tụ vào các "xung" dữ liệu thực tế: Với mỗi mẫu (h, r, t, τ_+) , lấy mẫu một thời điểm sai τ_- cực gần (ví dụ $\tau_- = \tau_+ \pm \delta$):

$$L_{time} = -\log \sigma(\phi(\tau_+) - \phi(\tau_-) - \gamma)$$

(γ là margin thời gian, khuyến nghị 0.1-0.5)

5.3 Regularizers

- **N3 Regularization:** Áp dụng cho h, t, r^H, r^T .
- **Amplitude Decay:** Phạt $\|A_{r,k}\|_2^2$ để tránh việc mô hình dùng quá nhiều năng lượng thời gian cho các quan hệ thực sự tĩnh.
- **σ Constraint:** Giới hạn σ không được quá lớn (tránh biến Gaussian thành hàm phẳng/tĩnh).

6. Khởi tạo Tham số (Initialization) - QUAN TRỌNG

- h, t, r^H, r^T : Khởi tạo Xavier/Glorot Normal.
- $\mu_{r,k}$: Rải đều (Uniform) trong đoạn $[0, 1]$. Ví dụ nếu $K = 10, \mu \in \{0.0, 0.1, \dots, 1.0\}$.
- $\sigma_{r,k}$: Khởi tạo nhỏ (ví dụ 0.05) để các xung ban đầu mang tính cục bộ cao.
- $A_{r,k}$: Khởi tạo rất nhỏ ($1e - 4$) để mô hình bắt đầu từ trạng thái Tĩnh, sau đó mới học cách Động hóa.

7. Quy trình Triển khai (Implementation Steps)

1. **Step 1: Normalize Time:** Ánh xạ `time_id` sang $\tau \in [0, 1]$ một cách tuyến tính.
2. **Step 2: Gaussian Encoder:** Tạo lớp `GaussianTimeEncoder` quản lý μ, σ, A .
3. **Step 3: Scoring:** Implement công thức cộng Residual $\Delta e_r(\tau)$ vào Entity Embeddings.
4. **Step 4: Hard Negative Sampler:** Viết hàm lấy mẫu τ_- lân cận cho L_{time} .
5. **Step 5: Training Loop:**
 - Epoch 1-5: Tắt L_{time} , chỉ học L_{fit} (Học thực thể tĩnh).
 - Epoch 6+: Bật L_{time} với trọng số tăng dần (Học xung thời gian).

8. Tại sao mô hình này vượt trội hơn TNTComplex?

- Khả năng nội suy: TNT không biết chuyện gì xảy ra giữa ngày T và $T + 1$. GE-PairRE tính toán được giá trị tại mốc $T.5$ nhờ hàm Gaussian liên tục.
- Khử nhiễu: Không giống như Sin/Cos. Gaussian không tạo ra "đuôi" dao động ở các vùng không có dữ liệu.