PREDICTED LINKS AS MESSAGES: SEPARATING FORMATION FROM PROPAGATION IN TEMPORAL GRAPHS

Anonymous authors
Paper under double-blind review

ABSTRACT

Temporal graph models propagate information along recorded interactions. This convention leaves a plausible future relation unavailable as context until it appears, even when the observed history already supports it. Admitting the relation early can connect useful representations, but it can also transmit an irrelevant state. A temporal learner must therefore decide which relation may form and how that relation should carry a message. We formalize this problem as *future-relation message admission*. Mechanism-Guided Graph Evolution (MGGE) uses historical formation evidence to freeze a sparse graph of schema-valid non-edges. A uniform residual channel exposes the graph without writing predicted relations into temporal memory. At the common $K = 8$ budget, paired 95% bootstrap confidence intervals for ΔMRR are positive on the corresponding evaluation populations of four real graphs, with gains ranging from 0.00105 to 0.01188 MRR. Equal-budget controls separate the effect of graph membership from edge count. Candidate-level analysis further reveals that formation rank does not provide a monotone ordering of message utility. On Wiki and GoogleLocal, two-sided exposure consistently exceeds source-only exposure across the tested budgets. A directional decomposition identifies the additional path through the destination representation. These results cast predicted links as candidate communication channels whose value depends on both admission and exposure.

1 INTRODUCTION

Temporal graph models support recommendation, communication analysis, and forecasting in systems whose relations arrive as timestamped events. Continuous-time learners differ in how they update memory, attend to temporal neighborhoods, or encode node pairs (Trivedi et al., 2019; Xu et al., 2020; Rossi et al., 2020; Yu et al., 2023; Zhang et al., 2025). These designs have made recorded history increasingly expressive. More demanding benchmarks also expose whether a model learns sequential structure rather than memorizing repeated interactions (Huang et al., 2023; Yi et al., 2025b). In standard observed-event-driven predictors, however, the propagation context is still constructed from recorded interactions.

Consider a user u who has interacted with several pages in one topical community. A candidate relation (u, v) may be plausible because it closes a typed structural pattern and points to an active destination. Exposing v as an auxiliary neighbor could change the score of a later query (u, w), even when (u, v) is not the query target. Under two-sided exposure, the same candidate can also change a later query (u', v) through the destination representation. It could help either query or inject an irrelevant state. Closure and activity describe why (u, v) may form; they do not determine how its representation will affect later predictions. *Formation plausibility is not message utility.*

This missing decision lies between link prediction and message passing. Link prediction scores a candidate pair when that pair is the query target. It does not decide whether the pair should alter other predictions before it occurs. Graph rewiring changes computational connectivity, but an added connection need not describe a relation that could plausibly appear in the system (Franceschi et al., 2019; Zhang et al., 2023; Barbero et al., 2024; Petrović et al., 2024; Linkerhägner et al., 2025). Network-formation models explain likely edge choices through closure, popularity, or community

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

structure (Overgoor et al., 2019; Gupta & Porter, 2022). Their formation likelihood does not measure the effect of an early message on a temporal learner. The unresolved question is which historically plausible relations should become computational context before they are observed.

We call this decision *future-relation message admission*. It separates two estimands. Conditional formation plausibility asks which legal destination is likely when a source initiates a new relation. Propagation utility asks how later predictions change after a frozen set of relations enters the message graph. We estimate the first quantity from historical choices and measure the second through a disjoint fit/support/query intervention. The resulting timeline fixes formation evidence before graph construction and fixes the auxiliary graph before query evaluation. Section 3 defines the two evaluation populations used to isolate indirect and full-population effects.

Mechanism-Guided Graph Evolution (MGGE) instantiates this decision with an auditable admission rule and a separate propagation intervention. Historical formation evidence ranks schema-valid non-edges, and a fixed budget converts the ranking into a frozen directed admission set. The temporal backbone consumes this graph through a uniform static residual. Formation scores determine membership, while a separate exposure rule determines which endpoints receive messages. Predicted relations remain outside temporal memory. Rebuilding the auxiliary graph at each evaluation episode lets the same admitted relations be studied under source-only and two-sided exposure.

The experiments yield two findings. First, sparse future-relation graphs improve later first-relation ranking on Wiki, Enron, UCI, and GoogleLocal at the common $K = 8$ budget. Equal-budget controls show that edge count alone does not explain the changes. Candidate-level utilities and controlled admission models also reveal that better formation fit need not produce a more useful message graph. Formation is a proposal signal, not a downstream utility score. Second, the way an admitted edge enters computation matters. Two-sided exposure exceeds source-only exposure at every tested budget on Wiki and GoogleLocal. The directional decomposition explains the added path: source-only exposure perturbs the selecting source, whereas two-sided exposure also lets that source state reach later queries through the admitted destination.

Contributions. Our main contributions are:

- **Problem.** We formulate future-relation message admission, separating the plausibility that a relation forms from its utility as an early message.
- **Intervention.** We introduce a history-measurable construction that separates formation-based membership from endpoint exposure, and derive the additional query path created by two-sided propagation.
- **Evidence.** Paired interventions on four real temporal graphs show when sparse future-relation messages improve first-relation ranking, distinguish graph membership from edge count, and identify endpoint exposure as a repeatable determinant of utility.

2 Related work

2.1 Temporal prediction from recorded history

Temporal predictors build message context from the recorded event stream. Recurrent state, attention, common-neighbor features, and MLP or mean-pooling encoders differ in how they represent that history (Trivedi et al., 2019; Xu et al., 2020; Rossi et al., 2020; Yu et al., 2023; Zhang et al., 2025; Cong et al., 2023). CRAFT uses destination-conditioned cross-attention to model seen and unseen future links (Yi et al., 2025a). TGM provides a unified library for temporal-graph learning, whereas Temporal Graph Thumbnail summarizes global temporal evolution with a compact skeleton (Chmura et al., 2026; Shi et al., 2026). These works enrich the representation or infrastructure used to model recorded history and score queried pairs. Future-relation message admission instead asks when an absent relation should modify the context of other queries before its first observation.

Chronological first-relation evaluation isolates this setting. TGB standardizes temporal link prediction, while TGB-Seq stresses sequential dynamics and unseen-edge generalization (Huang et al., 2023; Yi et al., 2025b). EdgeBank captures repeated-edge memorization; PopTrack shows that recent-node popularity is a strong baseline, and later studies document sensitivity to negative sampling (Poursafaei et al., 2022; Daniluk & Dabrowski, 2023; Romero et al., 2024). We compare paired interventions with fixed negatives on first-relation queries.

2.2 TASK-DRIVEN TOPOLOGY ADAPTATION

Graph structure learning and rewiring treat new connectivity as a task-dependent computational resource. TGSL learns time-aware candidate edges jointly with a temporal predictor, whereas TGR uses expander propagation to connect temporally distant nodes (Zhang et al., 2023; Petrović et al., 2024). Other work studies locality-aware rewiring, joint spectral rewiring and feature denoising, and the effectiveness and hyperparameter sensitivity of curvature-based rewiring (Barbero et al., 2024; Linkerhägner et al., 2025; Tori et al., 2025). TNCN instead enriches pairwise prediction with temporal common-neighbor representations (Zhang et al., 2025). Future-relation message admission assigns a different semantics to an added edge. Each admitted pair is a schema-valid absent relation supported by historical formation evidence. It enters a frozen, memory-external channel whose membership and endpoint exposure can be varied independently. This separation makes it possible to measure whether a plausible future relation is useful as a message, rather than absorbing admission and propagation into one topology objective.

2.3 FORMATION EVIDENCE FOR ADMISSION

Classical link-prediction work uses structural proximity, including common-neighbor signals, to identify candidate relations, while community-detection work characterizes mesoscopic graph structure (Adamic & Adar, 2003; Liben-Nowell & Kleinberg, 2007; Newman, 2006). For edge formation, conditional discrete choice combines mechanisms such as preferential attachment and triadic closure in destination-selection models (Overgoor et al., 2019); mixed logit allows these preferences to vary across nodes (Gupta & Porter, 2022). MGGE uses historical formation evidence only to determine membership in a directed admission set. A separate, score-free exposure operator determines which endpoints receive the static residual message, so formation fit and downstream message utility remain distinct.

3 FUTURE-RELATION MESSAGE ADMISSION

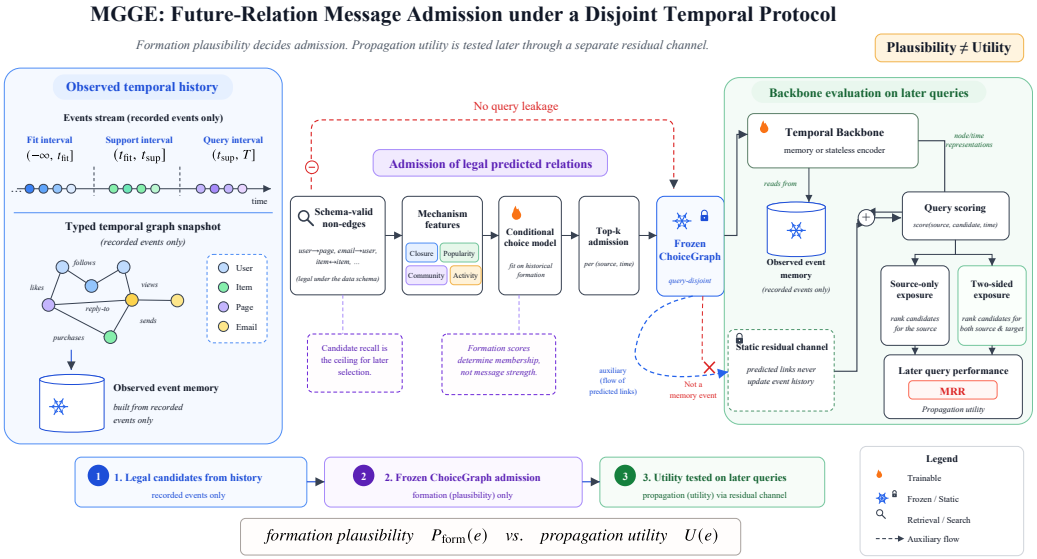


Figure 1: Future-relation message admission separates two decisions. Historical evidence freezes a directed admission set A_t before later queries. A separate exposure rule determines how those relations enter the static message graph B_t. Only observed support events update temporal memory.

Table 1: Core notation. Every time-indexed object used to construct the auxiliary graph is computed before the query interval.

Symbol	Meaning
$\mathcal{H}_{<t}$	observed event history before cutoff t
$\mathcal{D}(u)$	schema-valid destination nodes for source u
$\mathcal{U}_t(u)$	legal non-neighbors of u at t
$Z_t(u)$	indicator that u initiates a new relation after cutoff t
$\mathcal{S}_t(u)$	sampled alternatives used for choice estimation
$\mathbf{z}_{uv,t}$	standardized historical mechanism features
$s_{uv,t}$	frozen relative formation score
$\mathcal{P}_t(u)$	retrieved candidate pool, equal to $\mathcal{U}_t(u)$ when feasible
$\mathcal{C}_t(u)$	top-K admitted destinations for source u
A_t	directed set of admitted source–destination relations
B_t	static message-exposure graph derived from A_t
$\lambda \geq 0$	static residual strength
Q	query relations strictly after support

3.1 Decision problem and temporal boundary

Each evaluation episode follows a fixed fit/support/query timeline. The fit interval estimates formation from historical first choices. Formation parameters and the pre-support snapshot are fixed before the support interval; the observed support interval then identifies support-active sources, for which we construct and freeze the auxiliary graph before replaying support events. The query interval evaluates every graph condition with the same checkpoint, negatives, and edge budget. The backbone checkpoint is trained from observed events only; the auxiliary residual is applied after the temporal representation is computed for evaluation. Predicted relations can carry residual messages, but they never become events in temporal memory.

Let $\mathcal{H}_{<t}$ be the filtration generated by observed events before time t, and let $G_{<t} = (\mathcal{V}, E_{<t})$ be its structural projection. The interval endpoints satisfy $t_{\text{fit}} < t_{\text{sup}} < t_{\text{qry}}$. We write $Z_t(u) = 1$ when source u initiates a new relation after cutoff t. The destination-choice component conditions on this activation event rather than modeling it.

The schema-valid non-edge domain is

$$\mathcal{U}_t(u) = \mathcal{D}(u) \setminus \big(N_{<t}(u) \cup \{u\}\big). \tag{1}$$

For a bipartite interaction graph, $\mathcal{D}(u)$ contains only destination-type nodes and $N_{<t}(u)$ contains observed outgoing destinations. For a homogeneous directed graph, $\mathcal{D}(u)$ contains observed nodes other than u and $N_{<t}(u)$ again denotes outgoing destinations. The two target quantities are

$$F_t^{\text{dest}}(u, v) = \Pr[T_{uv} \in (t, t + \Delta] \mid Z_t(u) = 1, \mathcal{H}_{<t}], \tag{2}$$

$$U_t(B; Q) = M(f_\theta, \mathcal{H}, B; Q) - M(f_\theta, \mathcal{H}, 0; Q), \tag{3}$$

where T_{uv} is the first appearance time of relation (u, v) and M is a query metric such as mean reciprocal rank. The first quantity ranks destinations conditional on source activation. It does not model whether or when u becomes active. The second quantity measures a set-valued intervention on the backbone. For a candidate relation e, the corresponding marginal diagnostic relative to an exposure graph B is

$$U_t(e \mid B; Q) = M(f_\theta, \mathcal{H}, B \cup \phi(\{e\}); Q) - M(f_\theta, \mathcal{H}, B; Q). \tag{4}$$

We use formation evidence to constrain B but do not assume that F_t^{dest} is monotone in U_t. Appendix A gives a sufficient alignment condition. The experiments test whether that alignment holds in practice.

3.2 An auditable admission rule

The framework only requires an admission score computed from historical evidence. Our implementation uses an auditable mechanism vector for each legal pair,

$$\mathbf{x}_{uv,t} = \big[x_{uv,t}^{\text{clo}}, x_{v,t}^{\text{pop}}, x_{uv,t}^{\text{com}}, x_{v,t}^{\text{act}}, x_{v,t}^{\text{rec}}\big]^\top. \tag{5}$$

Closure is a common-neighbor count on homogeneous graphs and a typed length-three path count on bipartite graphs. The remaining features encode destination popularity, community compatibility, recent activity, and recency. Every component is a function of $\mathcal{H}_{<t}$.

Features are standardized on the fit data to obtain $\mathbf{z}_{uv,t}$. For an observed first relation (u, v), the estimation set contains the chosen destination and m alternatives sampled uniformly without replacement from legal historical non-neighbors,

$$\mathcal{S}_t(u) = \{v\} \cup \{w_1, \dots, w_m\}, \qquad w_i \sim \mathrm{Unif}(\mathcal{U}_t(u) \setminus \{v\}). \tag{6}$$

The default conditional multinomial logit model is

$$p_t(v \mid u, \mathcal{S}_t) = \frac{\exp(\boldsymbol{\beta}^\top \mathbf{z}_{uv,t})}{\sum_{w \in \mathcal{S}_t(u)} \exp(\boldsymbol{\beta}^\top \mathbf{z}_{uw,t})}. \tag{7}$$

A shared intercept is omitted because it cancels within every choice set. The signed coefficients describe relative formation odds rather than causal effects or message weights. Appendix B examines heterogeneous formation profiles with a latent-class sensitivity and tests whether the improved formation model changes downstream message utility.

When full enumeration is feasible, $s_{uv,t} = \boldsymbol{\beta}^\top \mathbf{z}_{uv,t}$ ranks every legal destination. Larger typed domains use a sparse history-only pool $\mathcal{P}_t(u) \subseteq \mathcal{U}_t(u)$ assembled from typed closure, community, popularity, activity, and recency candidates. Appendix B gives the retrieval implementation and its recall ceiling. The admitted neighborhood is

$$\mathcal{C}_t(u) = \mathrm{TopK}_{v \in \mathcal{P}_t(u)} \, s_{uv,t}. \tag{8}$$

If fewer than K legal candidates are available, all are retained. The selected pairs form the directed admission set

$$A_t = \{(u, v) : v \in \mathcal{C}_t(u)\}. \tag{9}$$

The history-only retrieval pool and the fixed budget K fully determine admission.

3.3 How an admitted relation carries a message

Membership in A_t does not specify how an admitted relation participates in computation. An exposure rule ϕ converts A_t into a static message graph $B_t = \phi(A_t)$. Source-only exposure sends the destination representation to the selecting source. Two-sided exposure also sends the source representation to the admitted destination. The second direction makes that source state available whenever the destination appears in a later query. We write $\mathcal{N}_{B_t}(i)$ for the incoming message neighborhood of node i.

Let $\mathbf{h}_i^t$ be the representation produced by the temporal backbone and let $d_i(B_t) = |\mathcal{N}_{B_t}(i)|$. The decoder receives

$$\widetilde{\mathbf{h}}_i^t = \mathbf{h}_i^t + \lambda \frac{1}{\max(1, d_i(B_t))} \sum_{j \in \mathcal{N}_{B_t}(i)} \mathbf{h}_j^t. \tag{10}$$

Formation scores decide membership in A_t but do not weight Equation 10. Relative choice calibration therefore cannot change propagation amplitude after membership is fixed. The residual is applied after the temporal representation is computed, and its pairs never update event memory. Degree normalization also keeps the worst-case perturbation norm independent of the number of admitted neighbors (Proposition 8).

All principal interventions use the pre-specified residual strength $\lambda = 0.20$; the appendix reports the λ and aggregation sensitivity grid. This fixes message amplitude while the experiments vary graph membership or endpoint exposure.

The exposure rule changes the set of queries reached by the same admitted pair, while also changing which endpoints receive residual updates. Let $J_q(\mathbf{h}_{a_q}, \mathbf{h}_{b_q})$ be a differentiable objective for a later source–destination query. Proposition 7 shows that the additional first-order contribution of two-sided exposure for admitted relation (u, v) is

$$\Delta J_{uv}^{\mathrm{two}} - \Delta J_{uv}^{\mathrm{source}} = \alpha_v \sum_{q : b_q = v} \langle \nabla_2 J_q, \mathbf{h}_u \rangle + O(\|\delta \mathbf{h}\|_2^2). \tag{11}$$

Table 2: Real-graph evaluation. Query counts are means after the common post-freeze target filter for five-seed runs and the fixed count for GoogleLocal.

Data	Nodes	Events	Queries	Role
Wiki	9,227	157,474	1,075	admission and exposure
Enron	370	125,235	353	admission and exposure
UCI	3,799	59,835	3,016	admission and exposure
GoogleLocal	473,580	1,913,967	25,194	typed scale and exposure

Source-only exposure affects queries whose selecting source is u. Equation 11 opens a second path through later queries whose destination is v. Destination reuse determines where this path is available, but does not by itself determine its sign or magnitude. Two-sided exposure therefore changes message reach and endpoint coverage, in addition to changing the number of directed residual updates. Formation odds contain neither reuse nor decoder sensitivity, so propagation direction must be evaluated after admission.

3.4 Disjoint utility evaluation

Candidate features are evaluated on the pre-support snapshot. Observed support events then identify the active sources for which admission lists are instantiated. Every auxiliary graph is constructed and frozen before support replay and before query labels are inspected. Query scoring uses the same negatives, backbone checkpoint, and realized edge budget. Temporal separation places all queries after support. For the central indirect-effect comparison, a post-freeze filter retains queries whose positive pair is absent from every maximum-budget frozen graph. Because nodes recur across intervals, this filter defines a conditional evaluation population. We also report a full-population positive-edge ablation. For each query, it removes the positive pair from that query's residual while all negative candidates retain the shared residual state. This diagnostic uses the positive identity during scoring; graph membership and temporal memory remain fixed.

The common subset isolates indirect message effects from direct insertion of the queried edge. Appendix A formalizes historical measurability, the retrieval ceiling, score invariance, and the directional decomposition.

4 Experiments

The experiments ask which history-supported but unobserved relations are useful as messages. RQ1 measures the effect of a frozen ChoiceGraph. RQ2a separates membership from edge count, while RQ2b holds membership fixed to study endpoint exposure. Candidate- and model-level diagnostics then compare formation fit with propagation utility. Appendix B contains the full budget sweeps and residual variants.

4.1 Experimental design

Data and task. Table 2 summarizes the four main graphs. Wiki and GoogleLocal are typed user–item systems. Enron and UCI contain directed interactions. We evaluate the first appearance of a legal relation because the auxiliary graph is frozen while that relation is still absent. Standard all-interaction Temporal Graph Benchmark evaluation also includes repeated events and therefore measures a different prediction population (Huang et al., 2023).

Disjoint intervention. Candidate construction uses fit and support history, and every graph is frozen before query labels are inspected. We evaluate a common indirect-effect subset whose positive pair is absent from every maximum-budget graph and a complementary full-population positive-edge ablation. All conditions replay the same observed support events and use the same checkpoint, query negatives, and realized edge count. Auxiliary relations remain in the static residual channel.

The reference backbone is a PyTorch Geometric implementation of the Temporal Graph Network (TGN), using `TGNMemory`, a `TransformerConv` embedding module, and a multilayer-perceptron link predictor. Wiki, Enron, and UCI use five training epochs. GoogleLocal uses one

Table 3: Frozen future-relation messages at the common reporting budget $K = 8$. The upper block uses the common indirect-effect subset with memory-based TGN. The lower block uses every eligible query with a query-local mask and a stateless Mixer-style encoder. Changes are paired within rows, and auxiliary relations remain outside event state.

Backbone	Data	Queries	Base	ChoiceGraph	ΔMRR [95% interval]
Common indirect-effect subset					
TGN	Wiki	1,075	0.05367	0.06555	**+0.01188 [0.01001, 0.01370]**
TGN	Enron	353	0.05498	0.05964	**+0.00466 [0.00053, 0.00971]**
TGN	UCI	3,016	0.09118	0.09223	**+0.00105 [0.00045, 0.00217]**
TGN	GoogleLocal	25,194	0.17027	0.18093	**+0.01066 [0.00474, 0.01718]**
Full population with query-local mask					
Mixer-style	Wiki	1,136	0.10925	0.11270	**+0.00346 [0.00202, 0.00503]**
Mixer-style	UCI	3,110	0.11007	0.11350	**+0.00343 [0.00213, 0.00494]**

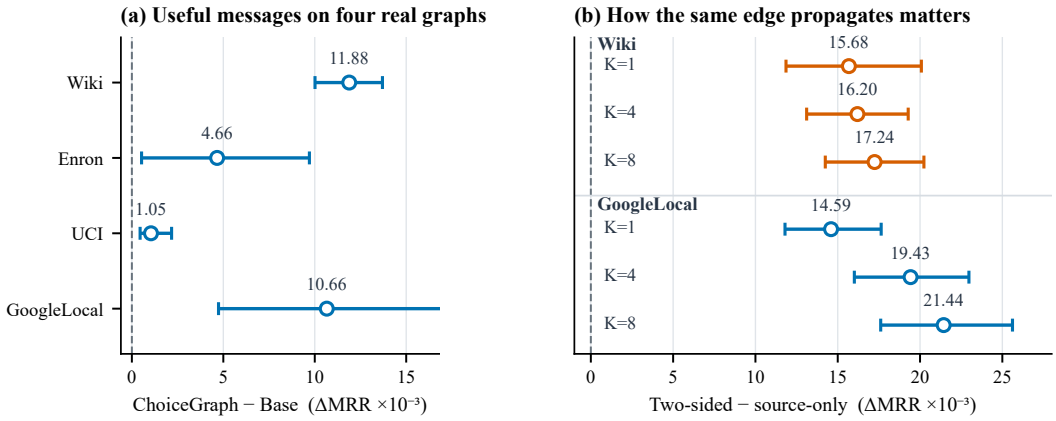


Figure 2: Core message-admission evidence. Panel (a) reports ChoiceGraph minus Base at $K = 8$ on four real graphs. Panel (b) holds the admitted graph fixed and reports two-sided minus source-only exposure at $K \in \{1, 4, 8\}$ on Wiki and GoogleLocal. Points are five-seed paired means. Whiskers are 95% seed-level bootstrap intervals.

complete epoch over 1,802,833 training interactions. We report mean reciprocal rank (MRR) over 100 negatives and percentile 95% bootstrap intervals over five seed-level differences.

Controls. Base uses no auxiliary graph. ChoiceGraph admits legal non-edges ranked by the conditional choice model. Random, degree-matched, popularity, and community controls change membership at the same realized edge budget; the future-support graph measures residual-channel capacity. Directional experiments instead fix membership and change endpoint exposure. We report $K = 8$ as the common cross-graph condition and provide every measured budget in Appendix B.

4.2 RQ1: Can an unobserved relation become a useful message?

At $K = 8$, the TGN comparison on the common indirect-effect subset is positive on Wiki, Enron, UCI, and GoogleLocal (Table 3). For scale, uniform ranking among one positive and 100 negatives has expected MRR $H_{101}/101 \approx 0.0515$; Base is 0.05367 on Wiki and 0.05498 on Enron. The effect varies with budget: Wiki and GoogleLocal are positive throughout the sweep, Enron strengthens as K grows, and UCI becomes positive at $K = 8$. The lower block transfers the intervention to a stateless Mixer-style encoder, with positive paired intervals on full-population Wiki and UCI queries.

Figure 4a places the common reporting budget in context. The differing curves show that message admission has a graph-dependent capacity regime. We use $K = 8$ as the shared cross-graph reporting condition.

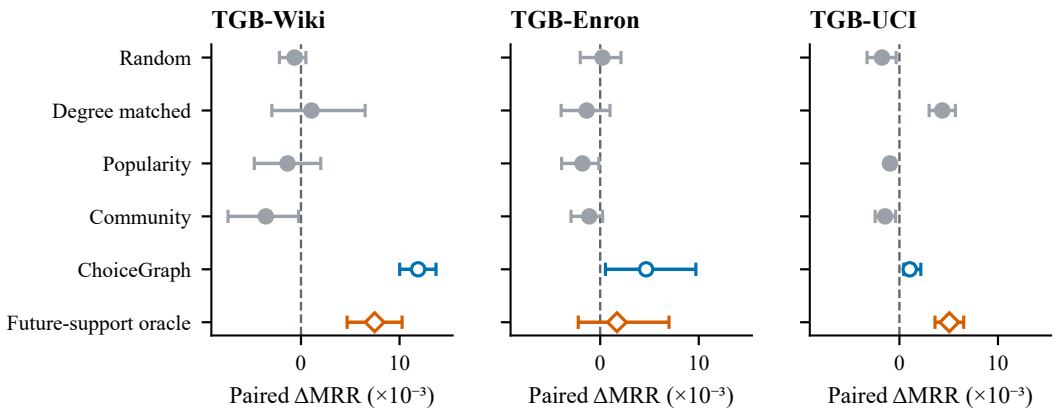


Figure 3: Equal-budget graph-membership controls at $K = 8$. Each point is the MRR of a frozen auxiliary graph minus Base on the same indirect-effect query subset. Whiskers are paired 95% bootstrap intervals over five seeds. The future-support condition measures residual-channel capacity.

GoogleLocal tests the same capability in a typed million-event domain. Its sparse history-only union scores 932 candidates per source, 0.35% of the legal destination domain, and recalls 20.04% of supported test destinations. Formation, retrieval, and graph freezing take 75.70 CPU seconds at 1,475 MiB peak RSS. The intervention retains 25,194 common queries per seed and improves MRR at $K = 8$.

4.3 RQ2A: Does graph membership matter beyond edge count?

Figure 3 fixes the checkpoint, edge budget, residual operator, negatives, and query set. Only graph membership changes. Random, degree-matched, popularity, and community graphs produce different utility patterns across Wiki, Enron, and UCI. ChoiceGraph is positive on all three. Adding K relations is therefore insufficient: the representations selected as new neighbors matter.

The Enron split also reveals how the indirect-effect population differs from the complete query set. The common subset retains 353 of 469 queries; the removed 24.73% have higher historical destination degree, with a retained-versus-removed standardized difference of approximately -0.75. On all 469 eligible queries, the positive-edge ablation changes MRR by 0.00329 [0.00016, 0.00675], providing a full-population sensitivity for the indirect-message result.

Learned admission controls. The matched study in Appendix B compares Choice, Neural, Recurrent, and Hybrid selectors with the same candidates, budget, checkpoint, residual, negatives, and query-local mask. Choice attains the highest mean MRR on Wiki and UCI, while Neural is highest on Enron. Selector flexibility alone therefore does not determine message utility.

4.4 RQ2B: How should the same relation enter computation?

Figure 2b freezes graph membership and changes only endpoint exposure. Every paired interval for two-sided minus source-only exposure is positive at $K \in \{1, 4, 8\}$ on Wiki and GoogleLocal. At $K = 8$, the gains are 0.01724 and 0.02144 MRR.

The two rules create different paths through later computation. Source-only exposure sends the selected destination state to source u, so it reaches queries whose source is u. Two-sided exposure also sends the source state to destination v, allowing that state to enter later queries whose destination is v. Equation 11 identifies this destination-mediated term, while degree normalization bounds each node's averaged residual. A Wiki control approximately matches directed message volume by giving source-only exposure twice the admission budget. The two-sided advantage remains positive at all three tested budgets (Table 4), linking the gain to endpoint access rather than update count.

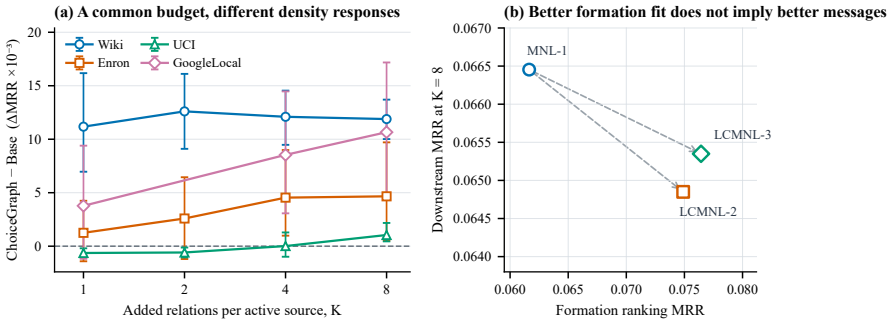


Figure 4: Message-admission diagnostics. Panel (a) reports ChoiceGraph minus Base across budgets; whiskers are paired 95% bootstrap intervals over five seeds. Panel (b) holds propagation fixed on Wiki: better formation ranking does not yield a detectable improvement in the message graph.

Table 4: Message-volume-matched direction control on Wiki. Each row compares two-sided exposure with K admitted relations against source-only exposure with $2K$ relations. Entries are paired MRR differences and 95% confidence intervals over five seeds.

K	Two-sided(K) − source-only$(2K)$, ΔMRR [95% CI]
1	0.00804 [0.00236, 0.01410]
4	0.01048 [0.00509, 0.01518]
8	0.01085 [0.00531, 0.01586]

4.5 Formation fit does not rank propagation utility

On Wiki, formation score is not a monotone proxy for message utility: across five seeds, Spearman correlation is -0.0409 [-0.0518, -0.0275]. The top four candidates have positive mean utility of 0.01251 [0.00476, 0.02010], followed by non-monotonic rank bands. Formation rank can propose low-budget members without ranking their message value.

Figure 4b tests the same distinction at the model level. Three held-out-selected classes raise Wiki formation MRR from 0.06164 to 0.07643 and Recall@64 from 27.25% to 31.08%, while their frozen graph reaches 0.06535 downstream MRR versus 0.06645 for one class (paired contrast -0.00110 [-0.00444, 0.00266]).

5 Limitations

Our estimand conditions on source activation and first-relation queries. The primary analysis measures indirect effects after excluding positive pairs contained in the frozen graph union; the full-population analysis is a query-local residual ablation. Architectural evidence covers a memory-based backbone and GraphMixer-style transfer on Wiki and UCI. The closure window, K, and λ remain schema-dependent, and choice coefficients describe associations rather than causal mechanisms.

6 Conclusion

Future relations are useful through the messages they enable, not their formation score alone. MGGE separates historical admission from message exposure in a frozen, memory-external channel. Equal-budget and message-volume controls show that both membership and exposure shape first-relation ranking. On Wiki and GoogleLocal, two-sided exposure adds a destination-mediated path and exceeds source-only exposure across the tested budgets.

REPRODUCIBILITY STATEMENT

All temporal splits are chronological. Formation and auxiliary-graph construction use fit/support history, and the central evaluation uses a common post-freeze indirect-effect subset. A complementary full-population analysis uses the positive identity for a query-local residual ablation while leaving the frozen graph and temporal memory unchanged. Candidate-generation budgets, random seeds, seed-level measurements, and paired confidence intervals are retained in machine-readable files. Oracle future edges appear only in a channel-capacity diagnostic. An anonymized code release is available at `https://anonymous.4open.science/r/mgge-anonymous-source-AEDE/`.

ETHICS STATEMENT

This work studies structural prediction on public or synthetic graph benchmarks. Predicted relations may be erroneous and should not be interpreted as verified social ties or deployed without domain-specific review. Any release will preserve the access and licensing terms of the source datasets.

AI USE STATEMENT

Generative AI tools assisted with literature organization, feedback on experiment and method design, code generation and inspection, experiment orchestration, result analysis, mathematical and conceptual exposition, figure preparation, and language editing. The authors reviewed and verified the research design, source code, experiments, numerical results, citations, claims, proofs, figures, and final manuscript, and take full responsibility for the work.

REFERENCES

Lada A. Adamic and Eytan Adar. Friends and neighbors on the web. *Social Networks*, 25(3): 211–230, 2003. doi: 10.1016/S0378-8733(03)00009-1.

Federico Barbero, Ameya Velingker, Amin Saberi, Michael Bronstein, and Francesco Di Giovanni. Locality-aware graph-rewiring in GNNs. In *International Conference on Learning Representations*, 2024. URL `https://openreview.net/forum?id=4Ua4hKiAJX`.

Jacob Chmura, Shenyang Huang, Tran Gia Bao Ngo, Ali Parviz, Farimah Poursafaei, Jure Leskovec, Michael Bronstein, Guillaume Rabusseau, Matthias Fey, and Reihaneh Rabbany. TGM: A modular and efficient library for machine learning on temporal graphs. In *International Conference on Learning Representations*, 2026. URL `https://proceedings.iclr.cc/paper_files/paper/2026/hash/d978cd64d598bbaf6a2fff1d6864d601-Abstract-Conference.html`.

Weilin Cong, Si Zhang, Jian Kang, Baichuan Yuan, Hao Wu, Xin Zhou, Hanghang Tong, and Mehrdad Mahdavi. Do we really need complicated model architectures for temporal networks? In *International Conference on Learning Representations*, 2023. URL `https://openreview.net/forum?id=ayPPc0SyLv1`.

Michal Daniluk and Jacek Dabrowski. Temporal graph models fail to capture global temporal dynamics. *arXiv preprint arXiv:2309.15730*, 2023. URL `https://arxiv.org/abs/2309.15730`.

Luca Franceschi, Mathias Niepert, Massimiliano Pontil, and Xiao He. Learning discrete structures for graph neural networks. In *International Conference on Machine Learning*, 2019. URL `https://arxiv.org/abs/1903.11960`.

Harsh Gupta and Mason A. Porter. Mixed logit models and network formation. *Journal of Complex Networks*, 10(6):cnac045, 2022. doi: 10.1093/comnet/cnac045.

Shenyang Huang, Farimah Poursafaei, Jacob Danovitch, Matthias Fey, Weihua Hu, Emanuele Rossi, Jure Leskovec, Michael Bronstein, Guillaume Rabusseau, and Reihaneh Rabbany. Temporal

graph benchmark for machine learning on temporal graphs. In *Advances in Neural Information Processing Systems: Datasets and Benchmarks Track*, 2023. URL https://arxiv.org/abs/2307.01026.

David Liben-Nowell and Jon Kleinberg. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, 58(7):1019–1031, 2007. doi: 10.1002/asi.20591.

Jonas Linkerhägner, Cheng Shi, and Ivan Dokmanić. Joint graph rewiring and feature denoising via spectral resonance. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/f463d31ed2fdd7b0ec585c041ec1baa8-Abstract-Conference.html.

M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006. doi: 10.1073/pnas.0601602103.

Jan Overgoor, Austin R. Benson, and Johan Ugander. Choosing to grow a graph: Modeling network formation as discrete choice. In *Proceedings of the 2019 World Wide Web Conference*, pp. 1409–1420, 2019. doi: 10.1145/3308558.3313662.

Katarina Petrović, Shenyang Huang, Farimah Poursafaei, and Petar Veličković. Temporal graph rewiring with expander graphs. *arXiv preprint arXiv:2406.02362*, 2024. URL https://arxiv.org/abs/2406.02362.

Farimah Poursafaei, Shenyang Huang, Kellin Pelrine, and Reihaneh Rabbany. Towards better evaluation for dynamic link prediction. In *Advances in Neural Information Processing Systems: Datasets and Benchmarks Track*, volume 35, 2022. URL https://openreview.net/forum?id=xHNzWHbklj.

Raphaël Romero, Maarten Buyl, Tijl De Bie, and Jefrey Lijffijt. Exploring the performance of continuous-time dynamic link prediction algorithms. *Applied Sciences*, 14(8):3516, 2024. doi: 10.3390/app14083516.

Emanuele Rossi, Ben Chamberlain, Fabrizio Frasca, Davide Eynard, Federico Monti, and Michael Bronstein. Temporal graph networks for deep learning on dynamic graphs. *arXiv preprint arXiv:2006.10637*, 2020. URL https://arxiv.org/abs/2006.10637.

Weining Shi, Zhisen Wen, Qinggang Zhang, Chentao Zhang, and Zhihong Zhang. Temporal graph thumbnail: Robust representation learning with global evolutionary skeleton. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/df29d63af05cb91d705cf06ba5945b9d-Abstract-Conference.html.

Floriano Tori, Vincent Holst, and Vincent Ginis. The effectiveness of curvature-based rewiring and the role of hyperparameters in GNNs revisited. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/ea7c7e886d48dd3a3a799819c8095012-Abstract-Conference.html.

Rakshit Trivedi, Mehrdad Farajtabar, Prasenjeet Biswal, and Hongyuan Zha. DyRep: Learning representations over dynamic graphs. In *International Conference on Learning Representations*, 2019. URL https://arxiv.org/abs/1803.04051.

Da Xu, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, and Kannan Achan. Inductive representation learning on temporal graphs. In *International Conference on Learning Representations*, 2020. URL https://arxiv.org/abs/2002.07962.

Lu Yi, Runlin Lei, Fengran Mo, Yanping Zheng, Zhewei Wei, and Yuhang Ye. Future link prediction without memory or aggregation. In *Advances in Neural Information Processing Systems*, volume 38, 2025a. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/036912a83bdbb1fd792baf6532f102d8-Abstract-Conference.html.

Lu Yi, Jie Peng, Yanping Zheng, Fengran Mo, Zhewei Wei, Yuhang Ye, Zixuan Yue, and Zengfeng Huang. TGB-Seq benchmark: Challenging temporal GNNs with complex sequential dynamics. In *International Conference on Learning Representations*, 2025b. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/db5ca61dbc08cf5143c05ad2d1b0b2ca-Abstract-Conference.html.

Le Yu, Leilei Sun, Bowen Du, and Weifeng Lv. Towards better dynamic graph learning: New architecture and unified library. In *Advances in Neural Information Processing Systems*, 2023. URL https://arxiv.org/abs/2303.13047.

Haozhen Zhang, Xueting Han, Xi Xiao, and Jing Bai. Time-aware graph structure learning via sequence prediction on temporal graphs. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pp. 3288–3297, 2023. doi: 10.1145/3583780.3615081.

Xiaohui Zhang, Yanbo Wang, Xiyuan Wang, and Muhan Zhang. Efficient neural common neighbor for temporal graph link prediction. In *Proceedings of the Fourth Learning on Graphs Conference*, volume 269 of *Proceedings of Machine Learning Research*, 2025. URL https://openreview.net/pdf?id=ffsDVGiKRm.