LATENT CAUSAL VOID: EXPLICIT MISSING-CONTEXT RECONSTRUCTION FOR MISINFORMATION DETECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Misinformation can remain locally coherent while misleading readers through omitted context. Detecting such cases requires comparing an article with contemporaneous reports, yet existing approaches often attach retrieved context as auxiliary evidence or describe hypothesized omission intent without specifying the missing content. We propose *Latent Causal Void* (LCV), a retrieval-guided detector that reconstructs this content as an explicit textual relation for cross-source reasoning. We formalize omission as an *environment-relative* discrepancy: information stated in a retrieved article but absent from the paired target sentence. LCV uses a frozen instruction-tuned large language model to generate a short missing-context description for each sentence–article pair, encodes these descriptions as relation embeddings in a heterograph over target sentences and context articles, and lets message passing incorporate the specific content missing from each sentence. On the bilingual benchmark of Sheng et al., LCV improves macro-F1 by 2.94 and 2.77 points on the English and Chinese splits over a matched reimplementation of the strongest omission-aware baseline. A human audit of 600 reconstructed relations finds 87% source support; among the 332 externally verifiable relations, 75% are independently corroborated. Controlled retrieval corruption across four families characterizes the detector's sensitivity to its information environment, while comparison with a strong non-graph model using the same reconstructed text isolates the contribution of graph reasoning. The content of an omission thus provides a useful unit of cross-source reasoning, making missing context available both for prediction and for direct inspection.

1 INTRODUCTION

Misinformation is often studied under the assumption that deception is expressed through what an article states, an assumption that has driven substantial progress in detecting fabricated claims, manipulated content, and misleading narratives Shu et al. (2017); Zhou and Zafarani (2020). Yet deception does not always arise from false statements. An article can be factually consistent at the sentence level while becoming misleading by selectively excluding contextual information necessary for interpreting the event. Such omission-based deception Carson (2010) poses a fundamentally different problem: the challenge is no longer identifying what is wrong in the text, but recovering what the text leaves unsaid.

Existing misinformation detectors largely reason over the presented content, exploiting stylistic patterns Xiao et al. (2024), knowledge-enhanced multimodal representations Liu et al. (2024); Wang et al. (2024); Zhu et al. (2025), persuasion-aware reasoning Modzelewski et al. (2025), external environment perception Sheng et al. (2022), or retrieval-based verification Yue et al. (2024); Wang et al. (2025) to determine whether the expressed claims conflict with available information. These approaches are effective when deception is explicitly encoded in the observed content, but provide limited mechanisms when misleadingness emerges from the relationship between what is stated and what is absent.

Omission-based deception remains comparatively underexplored because the missing information is not directly observable from the target article itself: an article may be locally coherent while selectively suppressing temporal, causal, or quantitative context required for a complete interpretation Wang et al. (2026); Sheng et al. (2022). Figure 1 illustrates this challenge, where contemporaneous reports reveal

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.

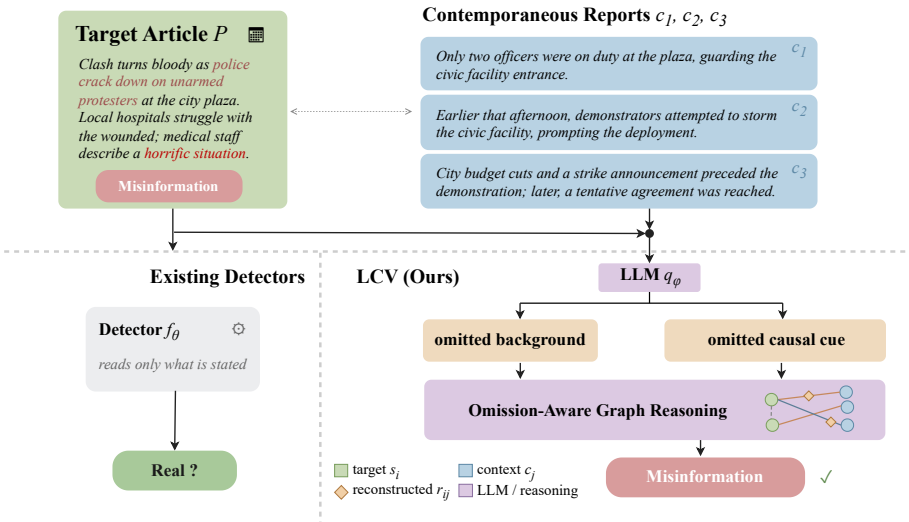


Figure 1: An omission-relevant misinformation case. The target article appears locally coherent, so a detector reasoning only over presented content may incorrectly regard it as real. Contemporaneous reports provide omitted background and causal information, including that only two officers were on duty and that protesters had attempted to storm government buildings. LCV reconstructs such missing content as textual cross-source relations and reasons over them in a heterograph. Following Definition 1, the reconstructed relations represent information supplied by the environment rather than independently verified facts.

additional circumstances that substantially change the interpretation of the event. Recovering these absent cross-source relationships, rather than merely checking individual statements, is therefore central to omission-aware detection; generation-side analyses further suggest that misinformation can emerge from real news through progressive omission of contextual details Liu et al. (2025a).

The key question is consequently not whether retrieved information exists, but how the discrepancy between a target and its information environment should be represented. Existing omission-aware approaches fall into three representational choices. *(i) Evidence attachment* retrieves related documents as auxiliary signals but leaves the target-context discrepancy implicit Sheng et al. (2022); Yue et al. (2024); Shukla et al. (2025); Xing et al. (2025). *(ii) Omission-intent description* generates explanations of why information may have been withheld and encodes them as relational attributes, capturing the existence of omission without specifying its content Wang et al. (2026). *(iii) Missing-content reconstruction* explicitly represents the information present in the retrieved environment but absent from the target as a textual relation. Only the third directly models the object that makes omission deceptive: the missing content itself.

Based on this observation, we propose *Latent Causal Void* (LCV), a retrieval-guided framework that reconstructs omitted content as explicit cross-source relations. Given a target article and temporally aligned context articles, LCV pairs target sentences with retrieved articles and uses a frozen instruction-tuned large language model to generate a concise proposition describing information present in the context but absent from the target. These propositions become relation embeddings in a heterogeneous graph over target sentences and context articles, where relation-aware message passing integrates the reconstructed missing context for prediction.

A central aspect of LCV is the formulation of omission as an *environment-relative* discrepancy rather than a claim against verified truth. The reconstructed relation specifies what the retrieved environment reports beyond the target sentence, without assuming that environment to be infallible. This matches the construction of contemporaneous information environments and makes the relations directly inspectable against both the source article and the target content.

Our main contributions are:

- We identify missing content as the representational unit of omission-based deception and propose reconstructing it as readable cross-source relation text, unifying existing omission-aware approaches.
- We formalize omission as an environment-relative discrepancy (Definition 1), capturing target-environment differences without treating retrieved context as verified truth.
- We introduce LCV, combining temporally aligned context retrieval, frozen-LLM missing-content reconstruction, and relation-aware heterograph reasoning.
- On bilingual benchmarks Sheng et al. (2022), LCV improves macro-F1 by $+2.94$/$+2.77$ over a matched reimplementation of OmiGraph, and we analyze the reconstructed relations through a 600-sample human audit, controlled retrieval corruption, representation-level comparisons, and architectural controls.

2 Related Work

2.1 Misinformation Detection from Presented Content

Most misinformation detection methods reason over information explicitly available in the target content. Early approaches focus on linguistic and stylistic signals Xiao et al. (2024), while subsequent studies incorporate richer semantic and contextual cues, including emotional patterns Zhang et al. (2021), persuasion-aware reasoning Modzelewski et al. (2025), knowledge-enhanced vision-language modeling Liu et al. (2024); Zhu et al. (2025), and manipulation analysis Wang et al. (2024); Zhang et al. (2025); recent efforts further explore causal explanation Chen et al. (2025a), multi-agent reasoning Li et al. (2025); Yan et al. (2025), contamination-aware evaluation Xu et al. (2025a); Xu and Yan (2025b), and retrieval-augmented inference Liu et al. (2025b).

A related direction extends detection beyond the target article by incorporating external information. NEP Sheng et al. (2022) models the surrounding news environment, RAFTS Yue et al. (2024) constructs contrastive arguments from retrieved evidence, RAV Shukla et al. (2025) integrates reconnaissance with verification, and MD-PCC Wang et al. (2025) identifies conflicts with commonsense knowledge. Other studies investigate complementary aspects, including multi-element retrieval Chen et al. (2025b), knowledge-graph reasoning Lourenço et al. (2026), temporal verification Barik et al. (2025), evidence attribution Xing et al. (2025), multilingual bias Vykopal et al. (2026), and adversarial robustness Schlichtkrull (2025). Despite their diversity, these approaches primarily treat external information as additional evidence for evaluating presented claims, leaving the discrepancy between what a target reports and what its environment contains implicitly represented rather than explicitly modeled.

2.2 Omission-Aware Detection

Omission-based deception introduces a different perspective: the challenge is not only determining whether stated information is incorrect, but identifying which absent contextual information changes interpretation Carson (2010); Greifeneder et al. (2020). Recent work has begun to explore this direction. OmiGraph Wang et al. (2026) constructs an omission-aware graph and uses an LLM to generate free-text descriptions for sentence–article pairs, encoding these descriptions as graph-edge attributes. Meanwhile, generation-side studies show that misinformation can emerge from authentic reports through progressive omission of contextual details Liu et al. (2025a).

Existing omission-aware methods differ primarily in how they represent the missing information. OmiGraph models an explanation of the potential omission process: its generated relation describes a hypothesized reason why information may have been withheld. In contrast, LCV models the missing content itself by generating a proposition that specifies information present in the retrieved environment but absent from the target. This distinction changes the object of reasoning: intent explanations characterize a possible mechanism of deception, whereas missing-content propositions directly expose the cross-source discrepancy that enables inspection against both the retrieved source and the target article.

LCV further formulates this discrepancy through an environment-relative omission definition (Definition 1), which specifies the relation between a target and its surrounding information environment

without requiring the environment to be treated as verified truth, enabling missing context to serve as an explicit relational object for downstream reasoning and human inspection. Table 8 summarizes the difference from existing omission-aware approaches.

3 METHOD

We present LCV as a retrieval-guided framework for omission-relevant misinformation detection. Figure 2 illustrates the overall pipeline, and Algorithm 1 summarizes the training procedure. The framework consists of five components: problem formulation, retrieval-guided graph construction, explicit missing-context reconstruction, omission-aware graph reasoning, and prediction with optimization.

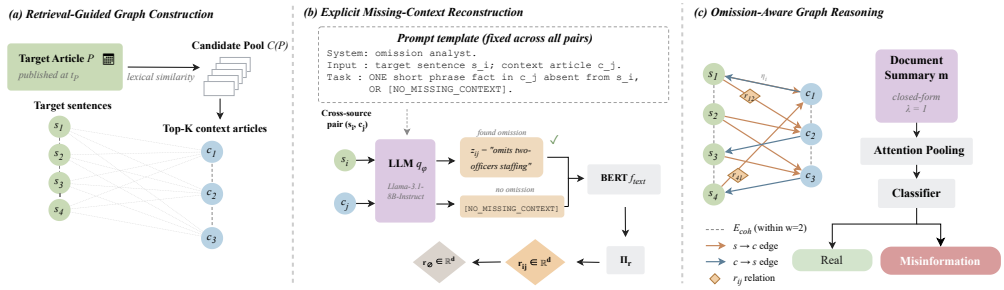


Figure 2: Overview of LCV. (a) Retrieval-guided graph construction selects the top-K temporally aligned context articles from $\mathcal{C}(P)$ within a Δ-day look-back window via TF–IDF cosine similarity, and assembles a heterograph over target sentence nodes $\mathcal{V}_s$ and context article nodes $\mathcal{V}_c$. (b) For each sentence–article pair $(s_i, c_j) \in E_{sc}$, a frozen LLM q_ϕ reconstructs a short missing-context description z_{ij}, which is encoded and projected into the cross-source relation embedding $\mathbf{r}_{ij}$. (c) Omission-aware graph reasoning propagates messages along coherence edges $\mathcal{E}_{\text{coh}}$ and bidirectional cross-source edges carrying $\mathbf{r}_{ij}$, combines them with a document-level summary $\mathbf{m}$, and predicts $y \in \{\text{real}, \text{misinfo}\}$.

3.1 PROBLEM FORMULATION

Let P denote a target candidate article published at time t_P with binary label $y \in \{0, 1\}$, where $y = 1$ denotes misinformation. The article is segmented into a sentence sequence $S(P) = \{s_1, \ldots, s_n\}$ with n bounded by a sentence budget. The candidate-context pool $\mathcal{C}(P)$ collects all articles published within a fixed look-back window $[t_P - \Delta, t_P)$ of temporal radius Δ. From $\mathcal{C}(P)$, a retrieval module selects the top-K relevant articles $\text{Ret}_K(P) = \{c_1, \ldots, c_K\}$. The cross-source pair set is $E_{sc} = \{(s_i, c_j) \mid s_i \in S(P),\ c_j \in \text{Ret}_K(P)\}$.

The prediction problem is to learn a classifier $f_\theta : (P, S(P), \text{Ret}_K(P), E_{sc}) \mapsto y$, where the central challenge is to recover cross-source content that is present in $\text{Ret}_K(P)$ but absent from $S(P)$, rather than merely aggregating observed text. The hyperparameters Δ, n, and K are fixed during training.

The recovered content is defined with respect to the information environment surrounding the target article, which motivates modeling the target–environment discrepancy directly.

Definition 1 (Environment-relative omission)**.** *Given a target sentence s_i from article P, a retrieved context article $c_j \in \text{Ret}_K(P)$, an* environment-relative omission *for the pair (s_i, c_j) is a proposition z that is entailed by c_j and is not entailed by s_i.*

Definition 1 defines a source-relative discrepancy: it specifies information surfaced by the context article but absent from the target sentence, without requiring the context article to represent verified truth. LCV therefore models the difference between target sentences and their retrieved environment rather than directly comparing the target with external facts. This matches the operation of the reconstruction module, which receives one sentence s_i and one context article c_j and identifies information stated by c_j but missing from s_i.

3.2 Retrieval-Guided Graph Construction

Following Sheng et al. Sheng et al. (2022), LCV ranks candidates by TF–IDF cosine similarity between the target article P and each $c \in \mathcal{C}(P)$ and selects the top-K as $\mathrm{Ret}_K(P)$. Target sentences and retrieved articles are assembled into a heterogeneous graph $\mathcal{G} = (\mathcal{V}_s, \mathcal{V}_c, \mathcal{E}_{\mathrm{coh}}, E_{sc}, \mathbf{R})$, where $\mathcal{V}_s = \{s_i\}_{i=1}^{n}$ are target-sentence nodes, $\mathcal{V}_c = \{c_j\}_{j=1}^{K}$ are context-article nodes, $\mathcal{E}_{\mathrm{coh}}$ are intra-target coherence edges, E_{sc} are cross-source sentence–article edges, and $\mathbf{R} = \{\mathbf{r}_{ij}\}$ collects the cross-source relation embeddings.

Let $f_{\text{text}}(\cdot)$ denote the shared frozen BERT Devlin et al. (2019) encoder used for sentences, articles, relation descriptions, and full documents, with output dimension $d_0 = 768$. Independent single-layer projections $\Pi_s, \Pi_c, \Pi_g, \Pi_r : \mathbb{R}^{d_0} \to \mathbb{R}^d$ map each representation type into its hidden space, giving $\mathbf{h}_i^{(0)} = \Pi_s f_{\text{text}}(s_i)$, $\mathbf{u}_j^{(0)} = \Pi_c f_{\text{text}}(c_j)$, and $\mathbf{g}^{(0)} = \Pi_g f_{\text{text}}(P)$ in $\mathbb{R}^d$.

3.3 Explicit Missing-Context Reconstruction

For each cross-source pair $(s_i, c_j) \in E_{sc}$, LCV reconstructs a short natural-language description $z_{ij} = q_\phi(s_i, c_j)$ of the content present in c_j but absent from s_i, following Definition 1. The instruction-tuned LLM q_ϕ remains frozen throughout training. The prompt requires either a concise phrase describing the missing content or the sentinel token `[NO_MISSING_CONTEXT]` when no such content is identified. It further instructs q_ϕ to report what the context article states rather than determine whether the statement itself is true; the complete prompt is provided in Appendix D, together with decoding and caching details.

The reconstructed text is encoded into the cross-source relation embedding

$$\mathbf{r}_{ij} = \Pi_r f_{\text{text}}(z_{ij}) \in \mathbb{R}^d,$$

and attached to both directions of the corresponding sentence–article edge. Sentinel outputs are mapped to a learnable null embedding $\mathbf{r}_\varnothing$, preserving graph connectivity while indicating the absence of informative missing content. Thus, retrieved context contributes to graph reasoning through explicit relation representations rather than document-level attachment alone. Figure 3 details the construction of these relations.

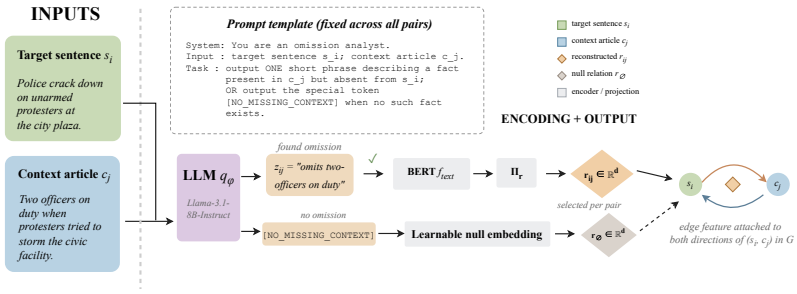


Figure 3: Construction of textual cross-source relations. For each sentence–article pair (s_i, c_j), the frozen LLM q_ϕ is queried with the fixed prompt template of Appendix D, which asks for one short phrase summarizing content present in c_j but absent from s_i, or the sentinel `[NO_MISSING_CONTEXT]` when no such content exists. Generated phrases z_{ij} are encoded by the shared BERT backbone f_{text} and projected by Π_r into the relation embedding $\mathbf{r}_{ij}$; sentinel outputs map to a null embedding $\mathbf{r}_\varnothing$.

3.4 Omission-Aware Graph Reasoning

The graph reasoner operates on $\mathcal{G}$ with a unified node hidden state $\mathbf{x}_a^{(\ell)} \in \mathbb{R}^d$ for each node $a \in \mathcal{V}_s \cup \mathcal{V}_c$ at layer ℓ, initialized as $\mathbf{x}_a^{(0)} = \mathbf{h}_i^{(0)}$ for $a = s_i$ and $\mathbf{x}_a^{(0)} = \mathbf{u}_j^{(0)}$ for $a = c_j$. The relation-type indicator $\rho(a, b) \in \{\mathrm{coh}, s \to c, c \to s\}$ gives the edge type between connected nodes; Figure 5 of Appendix C visualizes the reasoning process. efsec:recon.

Local sentence coherence. Target sentences are connected through a local window graph. With window size w, the coherence edge set is $\mathcal{E}_{\text{coh}} = \{(s_i, s_k) \mid 0 < |i-k| \le w\}$. For each coherence edge, LCV computes a difference-aware compatibility score:

$$\kappa_{\text{coh}}(\mathbf{x}_a^{(\ell)}, \mathbf{x}_b^{(\ell)}) = \exp\left(-\left\|\mathbf{W}_{\text{coh}}(\mathbf{x}_a^{(\ell)} - \mathbf{x}_b^{(\ell)})\right\|_2^2\right), \tag{1}$$

where $\mathbf{W}_{\text{coh}} \in \mathbb{R}^{d\times d}$ is a learnable projection. The coherence score is incorporated into the unified attention formulation as $\psi_{ab}^{(\ell)} = \log \kappa_{\text{coh}}(\mathbf{x}_a^{(\ell)}, \mathbf{x}_b^{(\ell)})$ for edges with $\rho(a,b) = \text{coh}$.

Relation-aware cross-source attention. For cross-source edges, the attention score jointly considers node states and reconstructed relation embeddings:

$$\psi_{ab}^{(\ell)} = \langle \mathbf{x}_a^{(\ell)}, \mathbf{Q}_{\rho(a,b)}^{(\ell)} \mathbf{x}_b^{(\ell)} \rangle + \langle \mathbf{r}_{ab}, \mathbf{T}_{\rho(a,b)}^{(\ell)} \mathbf{r}_{ab} \rangle, \tag{2}$$

where $\mathbf{Q}_\rho^{(\ell)}, \mathbf{T}_\rho^{(\ell)} \in \mathbb{R}^{d\times d}$ are learnable relation-specific bilinear forms. The first term captures node compatibility, while the second incorporates the semantic information carried by the reconstructed relation $\mathbf{r}_{ab}$.

The attention weight and node update are

$$\pi_{ab}^{(\ell)} = \frac{\exp(\psi_{ab}^{(\ell)})}{\sum_{b' \in \mathcal{N}(a)} \exp(\psi_{ab'}^{(\ell)})}, \qquad \mathbf{x}_a^{(\ell+1)} = \sigma\left(\sum_{b \in \mathcal{N}(a)} \pi_{ab}^{(\ell)} \mathbf{W}_{\rho(a,b)}^{(\ell)} \mathbf{x}_b^{(\ell)}\right), \tag{3}$$

where $\mathcal{N}(a)$ denotes the neighbors of node a in $\mathcal{G}$, $\mathbf{W}_\rho^{(\ell)} \in \mathbb{R}^{d\times d}$ is the relation-specific message transform, and $\sigma(\cdot)$ denotes the ReLU activation. After L propagation layers, the sentence states $\mathbf{h}_i^{(L)} \equiv \mathbf{x}_{s_i}^{(L)}$ encode both target-side coherence and cross-source relation information.

Document-aware global summary. LCV maintains a document-level summary $\mathbf{m} = (\sum_{i=1}^n \mathbf{h}_i^{(L)} + \lambda \mathbf{g}^{(0)})/(n+\lambda)$, where $\lambda \ge 0$ balances propagated sentence representations and the original document embedding. Each sentence representation is refined via a learned attention vector $\mathbf{a} \in \mathbb{R}^d$:

$$\tilde{\mathbf{h}}_i = \mathbf{h}_i^{(L)} + \eta_i \mathbf{m}, \qquad \eta_i = \frac{\exp(\langle \mathbf{a}, \mathbf{h}_i^{(L)} \rangle)}{\sum_{k=1}^n \exp(\langle \mathbf{a}, \mathbf{h}_k^{(L)} \rangle)}. \tag{4}$$

This aggregation incorporates document-level context while preserving sentence-level omission cues.

3.5 Prediction and Optimization

The target-side graph is summarized by attention pooling the refined sentence representations, $\mathbf{p} = \sum_{i=1}^n \beta_i \tilde{\mathbf{h}}_i$, where $\beta_i = \text{softmax}_i(\langle \mathbf{b}, \tilde{\mathbf{h}}_i \rangle)$ and $\mathbf{b} \in \mathbb{R}^d$ is a learnable pooling vector. The label distribution is $\Pr(y \mid P, \text{Ret}_K(P)) = \text{softmax}(\mathbf{W}_o \mathbf{p} + \mathbf{b}_o)$. Retrieval, LLM generation, and BERT encoding are precomputed and cached; during training we optimize the projection layers, graph reasoning parameters, global aggregation module, and classifier using the cross-entropy loss $\mathcal{L}_{\text{cls}} = \mathbb{E}_{(P,y)\sim\mathcal{D}}[-\log \Pr(y \mid P, \text{Ret}_K(P))]$.

4 Experiments

4.1 Datasets and Evaluation Protocol

Following Sheng et al. Sheng et al. (2022), we use one English dataset whose targets are Twitter posts and fact-checked claims, and one Chinese dataset whose targets are Weibo posts. Both are paired

with contemporaneous mainstream-news corpora, published within a seven-day window, that serve as the candidate-context pool $\mathcal{C}(P)$. We report macro-F1, accuracy, and per-class F1, following prior omission-aware comparisons.

One property of these benchmarks matters: targets are short (English averages 1.38 sentences), so most graphs are effectively bipartite with few coherence edges and the English split primarily tests cross-source relation modeling. The English split has 3900/1294/1289 train/val/test targets with 4.13 cross-source edges on average; Table 7 of Appendix H gives full statistics and Table 15 stratifies by target length.

4.2 Implementation Details

LCV retrieves the top-$K=3$ context articles within a $\Delta=7$-day look-back window using TF–IDF cosine similarity; targets are segmented into at most $n=10$ sentences. Propagation depth $L=2$, coherence window $w=2$, global mixing coefficient $\lambda=1$. Sentences, articles, and relation text are encoded with a shared frozen BERT Devlin et al. (2019) backbone (`bert-base-uncased` / `bert-base-chinese`; $d_0=768$, graph hidden $d=256$). Missing-context propositions are generated offline with `Llama-3.1-8B-Instruct` Dubey et al. (2024) using deterministic greedy decoding and a 16-token budget; the verbatim prompt is in Appendix D. We train with Adam (lr $=2\times10^{-5}$, batch 16, dropout 0.1, up to 30 epochs, patience 7, three fixed seeds).

4.3 Main Results

Table 1 reports the main comparison. LCV achieves the best macro-F1 on both datasets, improving over OmiGraph by $+2.94$/$+2.77$ (EN/ZH); paired t-tests over the three seeds are significant at $p<0.005$ and McNemar's gives $p<0.001$. The comparison uses a fixed retrieval pool and does not by itself verify that reconstruction recovered the right content, which the human audit addresses.

Table 1: Main comparison on the English and Chinese datasets. Higher is better. All methods, including OmiGraph, are run locally under the matched protocol of Section 4.1 and reported as mean±standard deviation over three seeds $\{42, 1024, 2048\}$. $*$ indicates that the improvement of LCV over OmiGraph is significant at the 0.005 level under a paired t-test over matched per-seed runs; ‡ additionally indicates significance under McNemar's test on pooled test-set predictions ($p<0.001$). Per-class F1 and accuracy are in Appendix H.

		English		Chinese	
Category	Method	macF1	Acc	macF1	Acc
Content-only	BERT Devlin et al. (2019)	0.7196±.0032	0.7135±.0021	0.7851±.0014	0.7921±.0016
	DualEmo Zhang et al. (2021)	0.7194±.0024	0.7200±.0021	0.7958±.0033	0.8003±.0029
	MSynFD Xiao et al. (2024)	0.7317±.0018	0.7319±.0017	0.8054±.0052	0.8089±.0048
	LLM Dubey et al. (2024)	0.5556±.0002	0.5779±.0000	0.6992±.0001	0.7110±.0001
	PCoT Modzelewski et al. (2025)	0.6458±.0011	0.6509±.0001	0.8020±.0001	0.8041±.0002
External info.-aware	NEP Sheng et al. (2022)	0.7274±.0004	0.7278±.0005	0.8288±.0010	0.8311±.0012
	MD-PCC Wang et al. (2025)	0.7227±.0028	0.7243±.0031	0.8168±.0022	0.8205±.0026
	RAV Shukla et al. (2025)	0.7189±.0020	0.7197±.0019	0.7930±.0018	0.7980±.0011
	RAFTS Yue et al. (2024)	0.6114±.0005	0.6049±.0006	0.7427±.0003	0.7580±.0002
Omission-aware	OmiGraph Wang et al. (2026)	0.7557±.0007	0.7647±.0011	0.8585±.0072	0.8596±.0065
	LCV (ours)	**0.7851±.0031***‡	**0.8067±.0029***‡	**0.8862±.0025***‡	**0.8827±.0022***‡

4.4 Ablation Study

We isolate the main design choices with three ablations: *w/o Retrieved Context* disables retrieval; *Structural Edges Only* replaces textual relation embeddings $\mathbf{r}_{ij}$ with uniform learnable attributes; *w/o Global Summary* removes the document-level summary $\mathbf{m}$. Table 2 shows that removing retrieved context drops macro-F1 by ~ 4.3 / ~ 4.2 points (EN/ZH), replacing textual relations by ~ 2.8 / ~ 2.6 points, and removing the global summary by ~ 1.2 / ~ 1.2 points. Ablation (ii) removes all per-edge information; the text-versus-intent-description comparison is Table 6.

Table 2: Ablation study on the English and Chinese datasets. Higher is better. All rows are means over the same three seeds $\{42, 1024, 2048\}$ under the matched protocol; per-seed standard deviations are reported in Table 9 of Appendix H. Note that *Structural Edges Only* removes all per-edge information rather than only its textual form; the text-versus-category comparison is Table 6.

	English		Chinese	
Variant	Macro-F1	Acc.	Macro-F1	Acc.
Full LCV	**0.7851**	**0.8067**	**0.8862**	**0.8827**
w/o Retrieved Context	0.7421	0.7618	0.8443	0.8431
Structural Edges Only	0.7568	0.7789	0.8597	0.8574
w/o Global Summary	0.7729	0.7936	0.8746	0.8718

4.5 What Do the Reconstructed Relations Actually Contain?

Downstream accuracy cannot establish that the reconstruction step recovered the right content, so we audit the relations directly. We sample 300 non-null relations per language, stratified over target labels and retrieval ranks, and have each judged independently by at least two annotators with a third adjudicating disagreements (Appendix E). The audit separates four properties that prior work tends to conflate: *source support* asks whether every proposition in z_{ij} is entailed by c_j; *independent factual correctness* asks whether it can be corroborated outside the benchmark corpus, scored only where such corroboration is obtainable; *attribution correctness* asks whether it is ascribed to the right source or event; and *completeness* measures coverage of annotation-marked omission units. Figure 4 gives a worked example.

Table 3: Human audit of reconstructed relations. Source support measures faithfulness to the retrieved article; independent factual correctness measures external corroboration and is computed only on the corroborable subset. Hallucination is reported per output and may co-occur with partial support. Full per-language breakdown in Table 10, Appendix H.

Property	English	Chinese	Pooled
Relations audited / corroborable	300 / 180	300 / 152	600 / 332
Source support (%)	89	84	87
Attribution correctness (%)	91	85	88
Independent factual correctness (%)	77	73	75
Completeness (%)	68	62	65
Hallucination rate (%)	11	14	13
Krippendorff's α	0.71	0.66	0.68

The gap between source support (87%) and independent corroboration (75%) reflects the corpus: the generator is faithful to what the retrieved article states, but retrieved articles are not always independently verifiable. Completeness is weakest at 65%, bounded by the 16-token format, and the Chinese split scores lower throughout, likely because its targets are longer and Weibo posts are stylistically distant from mainstream coverage.

4.6 Robustness to a Degraded Information Environment

We corrupt a fraction ρ of the top-K retrieved articles under four families: *irrelevant* (article from an unrelated event), *contradictory* (article asserting the negation of a true-context claim), *unreliable* (low-credibility coverage of the same event), and *incomplete* (article truncated to remove relevant content). Relations are regenerated from corrupted inputs at inference on a model trained under clean retrieval; Appendix G details the construction and Table 11 reports the full grid.

At $\rho = 0.25$, macro-F1 drops by 0.3–1.6 (EN) and 0.3–1.8 (ZH); at $\rho = 0.75$ the range widens to 3.0–9.2 and 4.0–9.8. Contradictory corruption is most damaging (roughly $3\times$ truncation at $\rho = 0.75$), consistent with a model that reads relation *content* rather than mere relation presence. LCV still exceeds clean OmiGraph under three of four families at $\rho = 0.25$; with one of three retrieved articles unreliable it exceeds matched clean OmiGraph by $+1.71$/$+1.31$ (EN/ZH).

4.7 Is the Graph Doing the Work?

The reconstruction step and the graph are separable contributions, and conflating them would overstate the architecture. We therefore compare against a strong non-graph consumer of exactly the same reconstructed text: DirectConcat($T{+}C{+}Z$) concatenates the target, the retrieved context, and the reconstructed relation text and classifies the result with no message passing, while DirectConcat($T{+}C$) removes the relation text to isolate its contribution. Both use the identical retrieval output, encoder, token budget, optimizer, splits, and seeds. Table 12 of Appendix H reports all metrics: relation text adds $+2.93$/$+2.70$ macro-F1 (EN/ZH) over plain concatenation, and graph reasoning a further $+1.27$/$+1.79$. Reconstruction is thus the larger contribution, and DirectConcat($T{+}C{+}Z$) already exceeds matched OmiGraph on both languages without any graph pipeline.

4.8 Text or Category? Isolating the Representation

To test whether missing-content text confers accuracy advantages over a fixed-type edge, we define a ladder: E0 uses a uniform learnable attribute; E1 assigns each z_{ij} to one of $M = 6$ omission categories[1] and uses a learned category embedding; E2 adds generator confidence; E3 is the textual embedding used throughout; E4 concatenates E1 and E3. The E3$-$E1 gap measures how much accuracy is gained by replacing a type-only edge with its textual content; a small gap would indicate that the contribution rests on Definition 1 and auditability rather than accuracy. Table 6 reports gaps of $+1.62$ (EN) and $+1.38$ (ZH), both exceeding E1$-$E0, so textual content carries discriminative signal beyond a type label.

4.9 Generator Sensitivity and Cost

LCV is not tied to a single generator (Table 5): Qwen2.5-7B costs 0.6/0.2 macro-F1 (EN/ZH) but stays above OmiGraph, while Mistral-7B degrades further and is weakest on format compliance; generator ordering matches for reconstruction quality and downstream accuracy. Generation is precomputed and cached, so LCV classifies at 48 ms per target (6.8 GB peak GPU); one-time generation scales as $n \times K$, i.e. 4.13 calls on average for English and 30 worst case, with the full breakdown in Table 13.

4.10 Limitations

LCV inherits the reliability of its information environment. Relations are faithful to the retrieved article rather than independently verified, so a uniformly corrupted pool propagates into the reconstructed content; the corruption study bounds this degradation rather than removing it, and credibility-aware weighting of context sources is a natural but unimplemented extension. Benchmark targets are also short, so the coherence mechanism is exercised mainly on the Chinese split and its benefit on long-form articles remains untested. The reconstruction step further inherits the generator's competence, since format compliance and hallucination both worsen on the weakest generator tested, and the audit covers 600 relations. Source-grounded reconstruction is therefore best suited to triage rather than autonomous verdicts.

5 Conclusion

LCV models omission-based misinformation by reconstructing missing-content propositions as textual cross-source relations. Three findings stand out. First, the environment-relative formulation decouples the learning signal from the truth status of the retrieved environment: even at $\rho = 0.25$ unreliable corruption LCV exceeds matched clean OmiGraph by $+1.71$/$+1.31$. Second, missing-content propositions yield $+1.62$/$+1.38$ macro-F1 (EN/ZH) with the architecture fixed, and a non-graph consumer of the same text already exceeds the strongest omission-aware baseline, establishing reconstruction as the primary source of gain. Third, the relations are inspectable, with 87% source support and 75% corroboration among the externally verifiable subset. Source-grounded propositions are thus a useful and auditable representation for context-aware misinformation detection.

[1] Missing background, causal antecedent, consequence, quantitative qualifier, counter-evidence, and temporal update.

AI USE STATEMENT

The method uses a frozen instruction-tuned language model, Llama-3.1-8B-Instruct, to generate the missing-context relation text described above; this model is an explicit component of the proposed system, not a source of undisclosed evaluation labels. Generated relation text is produced offline with a fixed decoding and post-processing procedure. All reported results and analyses were reviewed by the authors, who take responsibility for the final content and claims.

ETHICS STATEMENT

This work uses public benchmark data and releases no new high-risk dataset. Annotators in the relation audit (Appendix E) judged text from public news corpora, were compensated above local market rate, and were informed of the research purpose.

The system's error modes are asymmetric: its characteristic false positive is a truthful report that diverges from mainstream coverage, i.e., minority-viewpoint or early-breaking reporting. It should be used as decision support with human verification, not as an autonomous moderation tool. The method also inherits biases in the retrieved environment, and the corruption experiments characterize the resulting attack surface.

REPRODUCIBILITY STATEMENT

All hyperparameters, data splits, encoder checkpoints, LLM decoding configuration, and evaluation protocol are specified in Sections 3.1–4.2. All runs use seeds $\{42, 1024, 2048\}$. Appendix D gives the verbatim prompt for both languages; Appendix E gives the annotation guidelines; Appendix F details the OmiGraph adaptation; Appendix G describes the corruption construction. Both benchmarks and the OmiGraph code are publicly available.

REFERENCES

Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

Xinyi Zhou and Reza Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Computing Surveys*, vol. 53, no. 5, Art. 109, pp. 1–40, 2020.

Rainer Greifeneder, Mariela E. Jaffé, Eryn J. Newman, and Norbert Schwarz, Eds., *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation*. Routledge, 2020.

Thomas L. Carson, *Lying and Deception: Theory and Practice*. Oxford University Press, 2010.

Liang Xiao, Qi Zhang, Chongyang Shi, Shoujin Wang, Usman Naseem, and Liang Hu, "MSynFD: Multi-hop syntax aware fake news detection," in *Proceedings of the ACM on Web Conference 2024*, pp. 4128–4137, 2024.

Xuannan Liu, Peipei Li, Huaibo Huang, Zekun Li, Xing Cui, Jiahao Liang, Lixiong Qin, Weihong Deng, and Zhaofeng He, "FKA-Owl: Advancing multimodal fake news detection through knowledge-augmented LVLMs," in *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 10154–10163, 2024.

Bing Wang, Shengsheng Wang, Changchun Li, Renchu Guan, and Ximing Li, "Harmfully manipulated images matter in multimodal misinformation detection," in *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 2262–2271, 2024.

Ye Zhu, Yunan Wang, and Zitong Yu, "Multimodal fake news detection: MFND dataset and shallow-deep multitask learning," in *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 8012–8020, 2025.

Arkadiusz Modzelewski, Witold Sosnowski, Tiziano Labruna, Adam Wierzbicki, and Giovanni Da San Martino, "PCoT: Persuasion-augmented chain of thought for detecting fake news and

social media disinformation," in *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 24959–24983, 2025.

Qiang Sheng, Juan Cao, Xueyao Zhang, Rundong Li, Danding Wang, and Yongchun Zhu, "Zoom out and observe: News environment perception for fake news detection," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 4543–4556, 2022.

Zhenrui Yue, Huimin Zeng, Lanyu Shang, Yifan Liu, Yang Zhang, and Dong Wang, "Retrieval augmented fact verification by synthesizing contrastive arguments," in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 10331–10343, 2024.

Bing Wang, Ximing Li, Changchun Li, Bingrui Zhao, Bo Fu, Renchu Guan, and Shengsheng Wang, "Robust misinformation detection by visiting potential commonsense conflict," in *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 7760–7768, 2025.

Zhengjia Wang, Danding Wang, Qiang Sheng, Jiaying Wu, and Juan Cao, "Reasoning about the unsaid: Misinformation detection with omission-aware graph inference," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 40, no. 2, pp. 1249–1257, 2026.

Yuhan Liu, Zirui Song, Juntian Zhang, Xiaoqing Zhang, Xiuying Chen, and Rui Yan, "The stepwise deception: Simulating the evolution from true news to fake news with LLM agents," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 26176–26192, 2025.

Satyam Shukla, Himanshu Dutta, and Pushpak Bhattacharyya, "Recon, answer, verify: Agents in search of truth," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pp. 2429–2448, 2025.

Rui Xing, Timothy Baldwin, and Jey Han Lau, "Evaluating evidence attribution in generated fact checking explanations," in *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pp. 5475–5496, 2025.

Xueyao Zhang, Juan Cao, Xirong Li, Qiang Sheng, Lei Zhong, and Kai Shu, "Mining dual emotion for fake news detection," in *Proceedings of the Web Conference 2021*, pp. 3465–3476, 2021.

Tianlin Zhang, En Yu, Yi Shao, and Jiande Sun, "Multimodal inverse attention network with intrinsic discriminant feature exploitation for fake news detection," in *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 7940–7948, 2025.

Lizhi Chen, Zhong Qian, Peifeng Li, and Qiaoming Zhu, "Disconfounding fake news video explanation with causal inference," in *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 4842–4850, 2025.

Hui Li, Ante Wang, Kunquan Li, Zhihao Wang, Liang Zhang, Delai Qiu, Qingsong Liu, and Jinsong Su, "A multi-agent framework with automated decision rule optimization for cross-domain misinformation detection," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 5709–5725, 2025.

Zehong Yan, Peng Qi, Wynne Hsu, and Mong-Li Lee, "TRUST-VL: An explainable news assistant for general multimodal misinformation detection," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 5588–5604, 2025.

Cheng Xu, Nan Yan, Shuhao Guan, Yuke Mei, and M-Tahar Kechadi, "SSA: Semantic contamination of LLM-driven fake news detection," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 14737–14751, 2025.

Cheng Xu and Nan Yan, "TripleFact: Defending data contamination in the evaluation of LLM-driven fake news detection," in *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 8808–8823, 2025.

Xueqin Chen, Xiaoyu Huang, Qiang Gao, Li Huang, Jiajing Yu, and Guisong Liu, "Birds of a feather: Enhancing multimodal fake news detection via multi-element retrieval," in *2025 IEEE 41st International Conference on Data Engineering (ICDE)*, 2025, doi: 10.1109/ICDE65448.2025.00270.

Vítor N. Lourenço, Aline Paes, Tillman Weyde, Audrey Depeige, and Mohnish Dubey, "KG-CRAFT: Knowledge graph-based contrastive reasoning with LLMs for enhancing automated fact-checking," in *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 6419–6439, 2026.

Anab Maulana Barik, Wynne Hsu, and Mong Li Lee, "ChronoFact: Timeline-based temporal fact verification," in *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pp. 8031–8039, 2025.

Ivan Vykopal, Antonia Karamolegkou, Jaroslav Kopčan, Qiwei Peng, Tomáš Javůrek, Michal Gregor, and Marián Šimko, "Investigating language and retrieval bias in multilingual previously fact-checked claim detection," in *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 5195–5221, 2026.

Michael Sejr Schlichtkrull, "Attacks by content: Automated fact-checking is an AI security issue," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 8550–8565, 2025.

Zhiwei Liu, Kailai Yang, Qianqian Xie, Christine de Kock, Sophia Ananiadou, and Eduard Hovy, "RAEmoLLM: Retrieval augmented LLMs for cross-domain misinformation detection using in-context learning based on emotional information," in *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 16508–16523, 2025.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186, 2019.

Abhimanyu Dubey et al., "The Llama 3 herd of models," *arXiv preprint arXiv:2407.21783v1*, 2024.

A Training Procedure

Algorithm 1 summarizes training. Retrieval and relation generation are performed once, offline, before the training loop, so the per-epoch cost is that of the graph reasoner alone.

Algorithm 1 Training LCV.

Require: Target (P, y), retrieval budget K, propagation depth L, mixing coefficient λ

1: Retrieve $\mathrm{Ret}_K(P)$ from $\mathcal{C}(P)$ via TF–IDF cosine similarity
2: Segment P into sentences $S(P)$ and form the cross-source pair set E_{sc}
3: Generate $z_{ij} = q_\phi(s_i, c_j)$ and encode nodes/relations $\{\mathbf{h}_i^{(0)}, \mathbf{u}_j^{(0)}, \mathbf{g}^{(0)}, \mathbf{r}_{ij}\}$
4: **for** $\ell = 0, \ldots, L-1$ **do**
5: Update all $\mathbf{x}_a^{(\ell+1)}$ via the omission-aware graph operator
6: **end for**
7: Compute the document summary $\mathbf{m}$, refine $\{\tilde{\mathbf{h}}_i\}$, pool to $\mathbf{p}$, and predict $\hat{y}$
8: Update parameters by descending $\nabla_\theta \mathcal{L}_{\text{cls}}$

B Qualitative Case Study

Figure 4 shows a selective-reporting case from the English dataset. The target's three sentences are locally coherent; retrieved context articles supply budget-crisis background, stalled union negotiations, and a later tentative agreement. Panel 3 shows the reconstructed relations as sentence–article edges; Panel 4 shows the prediction shift from `Real` to `Fake` once these relations are integrated. The relations record what the coverage states, not what is independently verified (Definition 1).

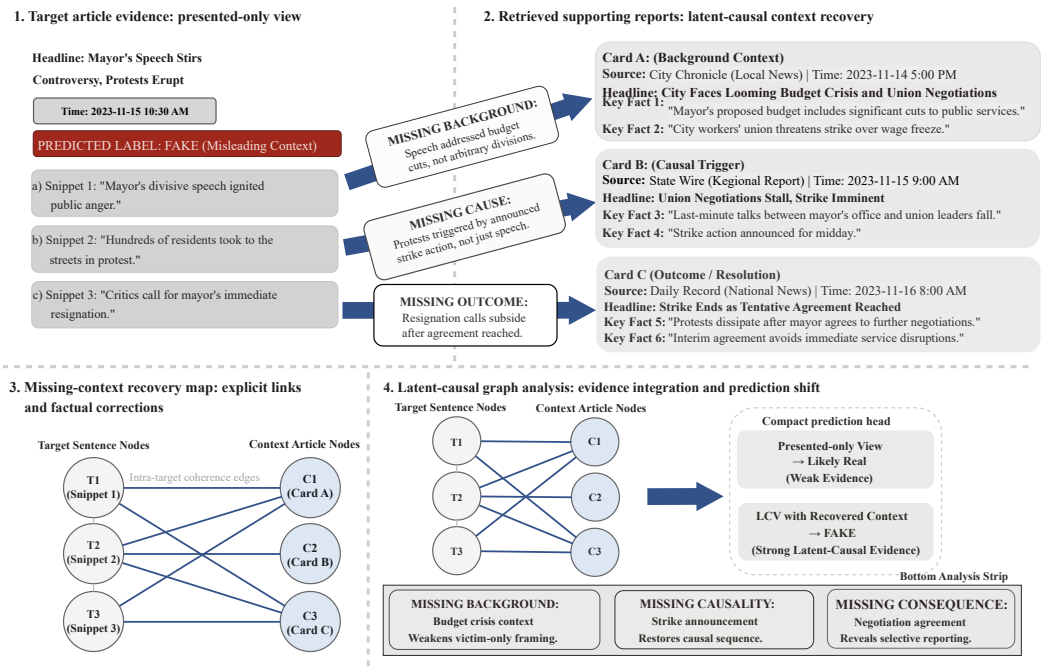


Figure 4: Qualitative LCV prediction on a selective-news-reporting case from the English dataset. Panel 1 contains the target article with three sentences $\{s_i\}$ that are locally coherent under a presented-only view. Panel 2 lists the retrieved context articles $\{c_j\}$ that supply the omitted background, causal trigger, and outcome. Panel 3 visualizes the reconstructed missing-context relations z_{ij} as dark sentence–article edges in the heterograph $\mathcal{G}$, with light edges denoting intra-target coherence in $\mathcal{E}_{\text{coh}}$. Panel 4 shows the same bipartite structure consumed by the omission-aware reasoner of Section 3.4: the compact prediction head moves from a weak `Real` decision under the presented-only view to a confident `Fake` decision once the recovered relations are integrated, with the bottom strip attributing the shift to missing background, missing causality, and missing consequence.

C Expanded Architecture Figures

The figure below expands the reasoning stage of Figure 2; the relation-construction stage is detailed in Figure 3.

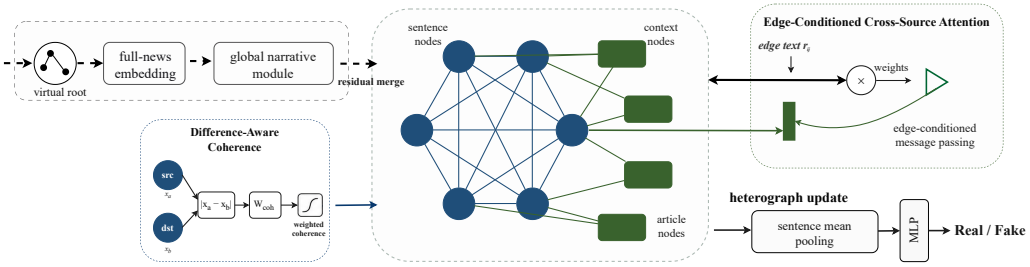


Figure 5: Graph-centric view of omission-aware reasoning. The *Difference-Aware Coherence* block (left) turns each coherence edge $(s_i, s_k) \in \mathcal{E}_{\text{coh}}$ into the weight $\kappa_{\text{coh}}(\mathbf{x}_a^{(\ell)}, \mathbf{x}_b^{(\ell)}) = \exp(-\|\mathbf{W}_{\text{coh}}(\mathbf{x}_a^{(\ell)} - \mathbf{x}_b^{(\ell)})\|_2^2)$ of Eq. equation 1. The *Edge-Conditioned Cross-Source Attention* block (right) fuses node states with the relation embedding $\mathbf{r}_{ij}$ via the bilinear score of Eq. equation 2, yielding edge-conditioned messages. Heterograph update (center) alternates the two message types over L layers; the full-news embedding $\mathbf{g}^{(0)}$ is residually merged with the propagated states into the global-narrative summary $\mathbf{m}$ of Eq. equation 4. Attention pooling of the refined states $\{\tilde{\mathbf{h}}_i\}$ with learned vector $\mathbf{b}$ (Sec. 3.5) produces the target representation $\mathbf{p}$, which feeds a linear classifier that predicts `Real`/`Fake`.

D Generation Prompt

The frozen generator q_ϕ is queried once per sentence–article pair with the template below. It reports what the context article *states*, not whether that statement is true. The Chinese template is a direct translation with identical structure and budgets.

English prompt.

```
[system]
You compare a target sentence against a context news article. You report
what the article states that the sentence does not. You do not judge
whether either text is true.

[user]
Context article (truncated to 256 tokens):
{c_j}

Target sentence (truncated to 32 tokens):
{s_i}

Output ONE short phrase naming a specific piece of information that the
context article states and the target sentence does not state. Do not
explain. Do not quote. Do not add commentary.
If the article states nothing that the sentence omits, output exactly:
[NO_MISSING_CONTEXT]

Answer:
```

The Chinese template mirrors this structure exactly; `[NO_MISSING_CONTEXT]` and the truncation budgets (256/32 tokens) are unchanged.

Decoding uses deterministic greedy decoding (`do_sample=False`) with a 16-token budget. Post-processing keeps only the first non-empty line, strips quotation marks and a leading `"Answer:"`,

normalizes variants of the sentinel to canonical `[NO_MISSING_CONTEXT]`, and truncates at the first sentence terminator. Outputs failing the one-phrase format are counted against the validity rate in Table 5. Generation is cached on disk keyed by pair hash and generator identifier.

E Human Annotation Protocol

Annotators. Six annotators participated, three per language. All hold at least a bachelor's degree, are native speakers of the language they annotated, and had prior experience with text annotation; two per language had prior fact-checking or journalism experience. Annotators were recruited from a university research pool, compensated above the local market rate for comparable work, informed of the research purpose, and free to withdraw. They did not see model predictions, target labels, or any indication of which system produced a relation.

Sampling. We sample 300 non-null relations per language from the test split, stratified jointly over the target label and over the retrieval rank of the context article so that neither real nor misinformation targets and neither high- nor low-ranked retrievals dominate. Relations whose output is the sentinel are excluded, since there is no proposition to audit. Each sampled item presents the annotator with the target sentence s_i, the full retrieved article c_j, and the generated relation z_{ij}.

Judgments. Annotators answer four questions per item, in this order, so that the source-relative judgment is made before any external search:

1. **Source support.** Is every proposition asserted by z_{ij} entailed by c_j? Answered on {fully supported, partially supported, unsupported}; only *fully supported* counts toward the reported percentage.
2. **Attribution correctness.** Is each proposition ascribed to the correct actor, source, or event? Items whose propositions carry no attribution are marked not-applicable and excluded from the denominator.
3. **Completeness.** Annotators first mark, independently of z_{ij}, which units of information in c_j are absent from s_i; completeness is the fraction of those units that z_{ij} covers.
4. **Independent factual correctness.** Can the proposition be corroborated by a source outside the benchmark corpus? Answered on {corroborated, contradicted, not corroborable}. Items marked *not corroborable* are excluded from the denominator, which is why the corroborable subset in Table 3 is smaller than the audited set. This question is answered last and with external search permitted.

Hallucination is recorded as a separate flag when z_{ij} asserts a proposition that appears in neither s_i nor c_j. It is reported per output and can co-occur with partial support, so it is not the complement of source support.

Adjudication and agreement. Every item is judged independently by two annotators. Disagreements on any of the four questions are resolved by a third annotator who sees both prior judgments without knowing their authors. Krippendorff's α is computed on the source-support judgment before adjudication, treating it as ordinal over the three levels, and is reported per language in Table 3. Agreement is lower on the Chinese split, consistent with the greater stylistic distance between Weibo targets and mainstream coverage.

Error taxonomy. Relations failing source support are additionally assigned to one of six categories, the same taxonomy used for the E1 rung of Table 6: missing background, missing causal antecedent, missing consequence, missing quantitative qualifier, missing counter-evidence, and missing temporal update. Category-level failure rates are reported in Table 16.

F OmiGraph Reimplementation

We use the authors' public implementation and change only the data-loading interface, so that the model consumes the same $\mathrm{Ret}_K(P)$ that LCV consumes rather than performing its own retrieval.

This is the single modification required to make the comparison matched, and it is the reason the numbers in Table 1 are the ones on which we run paired tests: with independent retrieval pools, per-seed pairing is not meaningful. All architectural hyperparameters, the omission-reasoning prompt, and the training schedule follow the strongest configuration reported by Wang et al. Wang et al. (2026). We train with the same three seeds and select checkpoints by validation macro-F1, as we do for every other system.

Table 4 reports our matched results next to the published point estimates. Readers comparing against the original paper should note that differences arise from the shared retrieval output and from our split construction, not from a change to the model.

Table 4: Matched local reimplementation of OmiGraph against the published configuration. Published values are point estimates without variance as reported in the original work; matched values are mean±std over seeds $\{42, 1024, 2048\}$ under our protocol.

Dataset	Source	Macro-F1	Accuracy
English	Published Wang et al. (2026)	0.7612	0.7653
	Matched reimplementation	0.7557±.0007	0.7647±.0011
Chinese	Published Wang et al. (2026)	0.8591	0.8603
	Matched reimplementation	0.8585±.0072	0.8596±.0065

G Retrieval-Corruption Construction

Corruption is applied at inference to a model trained under clean retrieval, which matches the deployment case where an environment degrades after a detector is fielded. For a target with retrieved set $\{c_1, \ldots, c_K\}$ we corrupt $\lceil \rho K \rceil$ articles, chosen from the lowest retrieval ranks first so that the highest-ranked evidence is the last to be lost. Corruption is applied to the article text before generation, so the reconstructed relations are regenerated from corrupted inputs rather than perturbed after the fact. The four families are constructed as follows.

Irrelevant replaces c_j with an article sampled uniformly from a different event cluster in the same corpus and the same time window, preserving length and register while removing topical relevance. *Contradictory* replaces c_j with an article about the same event in which a factual claim has been negated or its quantities altered, generated by rule-based edits over named entities, numbers, and polarity markers, then manually verified to read as fluent and self-consistent. *Unreliable* replaces c_j with coverage of the same event drawn from the lowest-credibility tier available in the corpus, preserving topicality while degrading provenance. *Incomplete* truncates c_j so that the span an annotator marked as supplying the omitted content is removed, while retaining the remainder of the article; this is the family most likely to yield sentinel outputs. Each condition is evaluated with the same three seeds, and the $\rho = 0$ column coincides by construction with the clean run in Table 1.

H Complete Result Tables

This appendix reports the full tables underlying every summary in the main text. All runs use seeds $\{42, 1024, 2048\}$ under the matched protocol of Section 4.1.

Table 5: Generator substitution, supporting Section 4.9. Valid output is the fraction of generations satisfying the one-phrase-or-sentinel format. Hallucination rate is reported as English / Chinese and is measured on the audited subset using the rules of Section 4.5.

Generator	Size	Valid (%)	EN macF1	ZH macF1	Halluc. (%)
Llama-3.1-8B-Instruct	8B	98.2	0.7851	0.8862	11 / 14
Qwen2.5-7B-Instruct	7B	96.7	0.7793	0.8841	13 / 15
Mistral-7B-Instruct-v0.3	7B	93.8	0.7614	0.8473	18 / 22

Table 6: Representation ladder, supporting Section 4.8. All rows share the identical architecture, retrieval output, encoder, optimizer, splits, and seeds, differing only in the edge representation $\mathbf{r}_{ij}$. E1 matches the representational commitment of a categorical omission-signal method within our architecture, so E3−E1 isolates the value of textual reconstruction.

	Edge representation $\mathbf{r}_{ij}$	English macF1	Chinese macF1
E0	Uniform learnable attribute	0.7568	0.8597
E1	Category embedding (M=6)	0.7689	0.8724
E2	Category embedding + confidence	0.7718	0.8751
E3	Textual embedding (LCV)	0.7851	0.8862
E4	Textual + category	0.7863	0.8874

Table 7: Processed split statistics, supporting Section 4. The average number of cross-source edges is the product of the average sentence count and the retrieval budget $K = 3$, and equals the average number of generator calls per target.

Dataset	Split	#Targets	Real	Misinfo.	Avg. #sent.	Avg. #context	Avg. #cross edges
English	Train	3900	1976	1924	1.378	3.000	4.134
	Val	1294	656	638	1.385	3.000	4.155
	Test	1289	661	628	1.463	3.000	4.389
Chinese	Train	4488	2304	2184	3.217	3.000	9.651
	Val	1496	768	728	3.194	3.000	9.582
	Test	1393	715	678	3.248	3.000	9.744

Table 8: Positioning of LCV against representative external-information and omission-aware detectors, supporting the discussion in Section 2.2. *Auditable relation* means the omission signal is emitted in a form a human annotator can check against its source. *Truth-agnostic* means the method does not require the retrieved context to be verified.

Property	NEP	RAFTS	OmiGraph	**LCV**
Uses external context	✓	✓	✓	✓
Models omission explicitly	—	—	✓	✓
Omission representation	—	—	category	**text**
Auditable relation	—	partial	—	✓
Truth-agnostic by construction	✓	—	not stated	✓

Table 9: Ablation study with per-seed variance, complementing Table 2.

	English		Chinese	
Variant	Macro-F1	Acc.	Macro-F1	Acc.
Full LCV	0.7851±.0031	0.8067±.0029	0.8862±.0025	0.8827±.0022
w/o Retrieved Context	0.7421±.0054	0.7618±.0048	0.8443±.0046	0.8431±.0041
Structural Edges Only	0.7568±.0045	0.7789±.0041	0.8597±.0039	0.8574±.0036
w/o Global Summary	0.7729±.0038	0.7936±.0034	0.8746±.0031	0.8718±.0028

Table 10: Complete human audit of reconstructed relations, complementing Table 3. Each audited relation is judged independently by at least two annotators with a third adjudicating disagreements. The contemporaneous corpus is not treated as verified ground truth, so independent factual correctness is scored only on the corroborable subset.

Dataset	Audited	Corrob. subset	Indep. correct (%)	Source support (%)	Attribution (%)	Completeness (%)	Halluc. (%)	Kripp. α
English	300	180	77	89	91	68	11	0.71
Chinese	300	152	73	84	85	62	14	0.66
Pooled	600	332	75	87	88	65	13	0.68

Table 11: Complete retrieval-corruption grid, supporting Section 4.6. Each cell reports Macro-F1 / Accuracy as mean±std over three seeds. ρ is the fraction of top-K retrieved articles corrupted; the $\rho=0$ column coincides with the clean run in Table 1. Construction of each corruption family is described in Appendix G.

Corruption	$\rho = 0.25$	$\rho = 0.50$	$\rho = 0.75$
English ($\rho = 0$: 0.7851±.0031 / 0.8067±.0029)			
Irrelevant	0.7782±.0044 / 0.7993±.0038	0.7631±.0058 / 0.7856±.0062	0.7418±.0079 / 0.7651±.0071
Contradictory	0.7693±.0052 / 0.7913±.0047	0.7362±.0073 / 0.7584±.0068	0.6934±.0098 / 0.7227±.0091
Unreliable	0.7728±.0049 / 0.7947±.0043	0.7497±.0064 / 0.7724±.0059	0.7213±.0082 / 0.7468±.0074
Incomplete	0.7821±.0036 / 0.8042±.0033	0.7726±.0047 / 0.7964±.0041	0.7547±.0061 / 0.7773±.0055
Chinese ($\rho = 0$: 0.8862±.0025 / 0.8827±.0022)			
Irrelevant	0.8784±.0033 / 0.8759±.0031	0.8617±.0051 / 0.8594±.0048	0.8354±.0072 / 0.8347±.0067
Contradictory	0.8683±.0047 / 0.8658±.0044	0.8347±.0068 / 0.8371±.0064	0.7878±.0093 / 0.7943±.0087
Unreliable	0.8716±.0044 / 0.8694±.0041	0.8452±.0062 / 0.8437±.0059	0.8134±.0083 / 0.8127±.0078
Incomplete	0.8834±.0028 / 0.8801±.0026	0.8678±.0042 / 0.8657±.0039	0.8467±.0059 / 0.8452±.0056

Table 12: Complete architecture controls, supporting Section 4.7. All controls use the same retrieval output, reconstructed relation text, text encoder, token budget, optimizer, splits, and seeds. DIRECTCONCAT variants receive no graph-specific features.

	English				Chinese			
Method	macF1	Acc	F1-real	F1-misinfo	macF1	Acc	F1-real	F1-misinfo
OmiGraph (matched)	0.7557±.0007	0.7647±.0011	0.7586±.0111	0.7528±.0093	0.8585±.0072	0.8596±.0065	0.8711±.0077	0.8460±.0061
DIRECTCONCAT(T+C)	0.7431±.0061	0.7642±.0057	0.7518±.0074	0.7344±.0079	0.8413±.0052	0.8449±.0048	0.8672±.0064	0.8154±.0068
DIRECTCONCAT(T+C+Z)	0.7724±.0048	0.7951±.0043	0.7783±.0059	0.7665±.0063	0.8683±.0038	0.8659±.0034	0.8831±.0047	0.8535±.0051
LCV (full heterograph)	0.7851±.0031	0.8067±.0029	0.7912±.0038	0.7790±.0041	0.8862±.0025	0.8827±.0022	0.9003±.0031	0.8721±.0034

Table 13: Computational cost, complementing Section 4.9. All rows measured on identical hardware, software, batch size, sequence lengths, and precision. NVIDIA RTX 3090 (24 GB), Intel Core i9-12900K, 64 GB RAM, float16, batch size 8, PyTorch 2.1.0, Transformers 4.38.2. The measured row uses the English split's empirical distribution of 4.13 pairs per target; the worst-case row uses the full budget n=10, K=3 and is the figure to budget against for long-form articles.

System	Cache setting	LLM calls / target	Generation time / target (s)	Classifier latency (ms)	Peak GPU (GB)	Throughput (targets/s)
LCV	Cached	0 (pre-cached)	2.3	48	6.8	20.8
LCV	Uncached, measured (EN)	4.13 avg	6.9	48	22.4	0.145
LCV	Uncached, worst case (n=10, K=3)	30	47.8	48	22.4	0.021
DIRECTCONCAT(T+C+Z)	N/A	0	0.8	42	4.2	23.8
BERT content-only	N/A	0	0.3	18	1.1	55.6

Table 14: Generation cost as a function of the sentence and retrieval budgets, holding all else fixed. The number of generator calls is $n \times K$ in the worst case and lower in practice because targets shorter than n contribute fewer pairs.

(n, K)	Worst-case calls	Generation time / target (s)
$(5, 1)$	5	8.2
$(10, 3)$	30	47.8
$(10, 5)$	50	79.4

Table 15: Results stratified by target length, supporting the discussion in Section 4.10. Targets are bucketed by sentence count; the English split is dominated by the single-sentence bucket, so the coherence term of Eq. equation 1 is largely inactive there.

	English		Chinese	
Bucket	#Test	Macro-F1	#Test	Macro-F1
$n = 1$	892	0.7814	347	0.8793
$n \in [2, 3]$	311	0.7926	682	0.8891
$n \geq 4$	86	0.8047	364	0.8942

Table 16: Distribution of relations without full source support across predicted missing-content categories. Percentages sum to 100% within each language; they reflect category composition among failing relations and should not be read as category-specific failure rates.

Omission category	English (%)	Chinese (%)
Missing background	31	28
Missing causal antecedent	22	24
Missing consequence	18	19
Missing quantitative qualifier	12	14
Missing counter-evidence	11	9
Missing temporal update	6	6