International Journal of Law, Ethics and Social Sciences

Article

Character Networks and Culprit Concealment in Classical Detective Fiction, 1868–1930

Xinyi Fan[1,*]

[1] International Division of the Experimental High School Attached to Beijing Normal University, Beijing, China

* Correspondence: Xinyi Fan, International Division of the Experimental High School Attached to Beijing Normal University, Beijing, China

Abstract: Detective fiction invites readers to use narrative detail as evidence, but the most visible character is not necessarily the culprit. This article examines twenty English-language detective novels published between 1868 and 1930 through character co-occurrence networks, applying digital humanities methodologies to classical narratology. The first 80% of each novel forms a pre-denouement reader network. Mention frequency, however, places a culprit among the top three candidates in 13 of 20 novels (65%; random expectation 31.4%, p = .002), with ten culprits ranked second or third (p = .003). Culprits often appear early, but so do other major characters. Within individual novels, early appearance does not distinguish culprits from other candidates (p = .918). In 13 of the 20 novels, the culprit ranks among the top three candidates in the change of narrative visibility between the investigation and the ending (p = .002). This result also holds under alternative measures. Furthermore, the novels reveal two distinct strategies for concealing a culprit: keeping the character relatively inconspicuous (under-exposure) or making the character familiar enough to escape suspicion (over-exposure). These empirical findings help explain how detective novels strategically guide readers' attention and manipulate narrative focus. Ultimately, the study demonstrates that while character networks illuminate structural patterns of concealment, computational metrics alone cannot definitively identify the culprit without qualitative contextualization.

Keywords: detective fiction; character networks; narratology; digital humanities; narrative visibility

1. Introduction

Detective novels constitute a distinctive literary genre characterized by their deliberate orchestration of narrative attention, where seemingly peripheral textual elements—such as fleeting character appearances, ambiguous dialogue tags, or incidental spatial references—may later assume decisive interpretive weight. A minor figure introduced in passing during an early chapter may ultimately be revealed as the central agent of criminal action, thereby challenging readers to maintain sustained vigilance over cumulative textual information. This structural principle is reinforced through long-standing conventions of fair-play storytelling, which emphasize equitable distribution of evidentiary clues across the narrative surface, enabling readers to engage in logical deduction on equal footing with the fictional investigator. Such conventions are not merely stylistic preferences but reflect deeper commitments to cognitive fairness, narrative transparency, and readerly agency within the genre's formal architecture [1].

Character network analysis offers a rigorous computational and structural framework for investigating how narrative attention is distributed, sustained, and modulated across fictional personae. By modeling characters as nodes and their interactions—as manifested through co-occurrence, dialogue adjacency, spatial proximity, or syntactic dependency—as edges, this approach transforms qualitative narrative phenomena into quantifiable relational structures. Early foundational work demonstrated

Received: 18 September 2026
Revised: 29 October 2026
Accepted: 10 November 2026
Published: 16 November 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

that macro-level patterns in these networks often correlate with genre-specific narrative functions, such as exposition, complication, or resolution [1, 2]. Subsequent methodological refinements have enabled more granular extraction of character relations from unstructured prose, incorporating advances in named entity recognition, coreference resolution, and discourse segmentation. However, significant variability remains in outcomes due to differences in preprocessing protocols—including thresholds for character merging, treatment of collective nouns, handling of pseudonyms or titles, and criteria for defining meaningful interaction—underscoring the need for standardized, reproducible pipelines grounded in literary-theoretical assumptions about character ontology and narrative function.

This study systematically investigates three interrelated research questions grounded in the structural logic of detective fiction [3]. First, it examines whether the culprit's position within the evolving character network—measured through centrality metrics such as degree, betweenness, eigenvector, and temporal persistence—exhibits statistically distinguishable patterns relative to baseline expectations, thereby enabling identification with significantly higher accuracy than chance alone (RQ1). Second, it conducts comparative topological analysis to determine how the culprit's network role diverges from that of other plausible suspects—not only in absolute centrality values but also in trajectory, clustering behavior, and structural equivalence across narrative phases (RQ2). Third, it interprets these quantitative distinctions qualitatively to classify concealment strategies employed by authors, distinguishing between obfuscation via narrative marginalization, strategic overexposure, functional misdirection, or relational camouflage—thereby contributing to a typology of narrative deception rooted in network dynamics rather than purely lexical or syntactic features (RQ3).

2. Theoretical Framework

Classical detective fiction is structured around a dual narrative architecture: one narrative recounts the commission of the crime itself, an event that occurs prior to the commencement of the investigative process, while the second narrative traces the methodical unfolding of the inquiry, during which evidence is gathered, interpretations are tested, and the underlying truth is progressively reconstructed. This structural duality establishes a temporal asymmetry between what has already transpired and what is being actively uncovered, thereby shaping both reader expectations and the logical scaffolding upon which narrative resolution depends. The separation of these two timelines is not merely formal but epistemological—what is known to the author and implied in the plot's design remains concealed from the reader until deliberate disclosure mechanisms are activated through dialogue, revelation, or narrative pivot [1].

To ensure methodological integrity and prevent circular inference, the analytical framework deliberately excludes the final segment of each novel when constructing the primary character interaction network. Specifically, the 'reader network' is derived exclusively from the initial eighty percent of the textual corpus, thereby simulating the informational constraints experienced by an actual reader progressing through the narrative without foreknowledge of the resolution [4]. This segmentation enables a temporally grounded assessment of character centrality, mitigating the risk that apparent predictive success arises not from meaningful structural patterns but from post-hoc exposure to the denouement. The remaining twenty percent—encompassing the climax, confession, and explanatory summation—is reserved for a distinct phase of analysis focused on dynamic shifts in relational prominence, particularly examining how attentional weight redistributes across characters as narrative certainty increases and interpretive ambiguity collapses. Such a staged approach supports robust validation of whether early-network features possess genuine diagnostic value independent of terminal textual cues.

Narrative conventions governing suspect presentation reflect longstanding tensions between fairness and misdirection. Certain compositional norms emphasize early introduction of the perpetrator to satisfy principles of equitable clue distribution, while

others advocate for sustained narrative visibility to maintain plausibility within the social fabric of the fictional world. Yet excessive prominence risks undermining the very mystery the genre seeks to sustain, transforming the culprit from a concealed agent into an overtly flagged figure. This paradox is rooted in the fundamental hermeneutic function of detective fiction: the text must simultaneously generate curiosity and defer resolution, sustaining interpretive tension through calibrated ambiguity. Effective storytelling thus requires careful modulation of referential density, syntactic framing, and discursive positioning—not only to obscure identity but also to preserve the coherence of alternative hypotheses throughout the investigative arc [5].

Four interrelated hypotheses guide the empirical investigation. The first explores temporal distribution, testing whether perpetrators consistently appear among the earliest introduced characters, thereby aligning with foundational expectations of narrative equity [2]. The second evaluates relative prominence, assessing whether culprits exhibit measurably higher centrality metrics—such as degree, betweenness, or eigenvector scores—compared to other major suspects across comparable narrative segments. The third hypothesis investigates dynamic trajectory, analyzing whether visibility undergoes statistically significant inflection precisely at the point of narrative resolution, suggesting a structural signature tied to revelation rather than gradual accumulation. The fourth hypothesis addresses generalizability, evaluating whether a composite profile—integrating multiple topological indicators—retains discriminative power when applied to previously unseen texts, thereby supporting claims of cross-novel applicability beyond training data boundaries.

Narrative perspective introduces critical complications for network-based character modeling, especially when reliance is placed solely on surface-level lexical markers such as proper nouns. In first-person narration, the narrator functions as both observer and participant, yet their self-reference typically occurs via pronouns ('I', 'me', 'my') rather than explicit naming, resulting in systematic underrepresentation if name frequency alone determines node inclusion or weight. This discrepancy becomes especially consequential in cases where the narrator occupies a pivotal role in the plot's causal logic—such as serving as the perpetrator—because standard co-occurrence or adjacency metrics fail to capture their pervasive discursive presence [6]. A canonical illustration is found in narratives where the narrator's identity is deliberately withheld until the conclusion; here, the absence of their name in early network construction does not reflect marginality but rather reflects the limitations of token-based extraction methods. Consequently, robust character network analysis must incorporate discourse-functional annotation, distinguishing between grammatical subjecthood, referential anchoring, and pragmatic agency to avoid distorting relational topology through superficial lexical criteria.

3. Data and Methods

This study conducts a systematic computational analysis of twenty canonical detective novels originally published between 1868 and 1930, all of which have entered the public domain under United States copyright law [7]. The selected corpus represents foundational works in the evolution of the detective fiction genre across multiple national traditions, including British, American, and early continental European influences. The full textual corpus comprises approximately 1.52 million words, ensuring sufficient lexical and structural diversity for robust network-based modeling. To support methodological rigor and generalizability, the dataset was deliberately partitioned into two functionally distinct subsets: eight novels published prior to 1920 were designated as the development set, used exclusively for feature engineering, model calibration, and profile construction; the remaining twelve novels—spanning diverse narrative structures, temporal settings, and authorial conventions—were strictly reserved for out-of-sample evaluation, thereby preventing any form of data leakage or overfitting during validation.

Character identification followed a rigorous, multi-stage manual annotation protocol. Each novel underwent line-by-line examination to identify all named individuals, with careful attention paid to variant spellings, honorifics, nicknames, pseudonyms, and

culturally embedded epithets that might refer to the same individual. For each pair of characters appearing within the same paragraph, a co-occurrence edge was instantiated in the character interaction graph [5]. Edge weights were computed cumulatively across the entire text, reflecting the total number of shared paragraphs—a quantitative proxy for narrative proximity and functional interdependence. Candidate selection applied a dual threshold: only characters mentioned at least five times within the first 80% of the narrative were retained, and no more than twelve candidates per novel were admitted to the final pool to ensure analytical tractability and comparability across heterogeneous texts. The definition of the central narrative agent—the figure whose actions precipitate the core plot resolution—was standardized across the corpus: in murder-centered narratives, this agent is the individual whose conduct directly causes the fatal outcome disclosed in the concluding revelation; in *The Moonstone*, where the primary narrative conflict centers on the unauthorized removal and subsequent recovery of a sacred artifact, the central agent is defined as the individual whose deliberate action initiates the chain of events culminating in the artifact's displacement from its original custodial context.

Analytical focus remained consistently anchored on narrative function rather than socioprofessional identity [4]. Consequently, occupational labels—such as 'butler', 'physician', or 'magistrate'—were treated solely as contextual descriptors and excluded from formal classification criteria. The principal investigator figure, recurring supporting agents who participate substantively in investigative procedures, and official authorities formally assigned to the case were all retained within the structural character network to preserve narrative topology and relational integrity; however, they were systematically excluded from the candidate pool to prevent circularity and ensure that predictive modeling targeted only those figures whose narrative roles remain ambiguous until the final resolution.

Specialized handling protocols were implemented for first-person narrative frameworks to mitigate pronoun-based overcounting artifacts [6]. Rather than enumerating every instance of personal pronouns, a paragraph-level attribution rule was adopted: in novels featuring a single, active narrator who also participates in the plotted events, a paragraph was assigned to that narrator if and only if it contained at least one unquoted occurrence of the following first-person lexical forms: 'I', 'me', 'my', 'myself', 'I'm', 'I'd', 'I've', or 'I'll'. Crucially, each qualifying paragraph contributed only a single count toward the narrator's overall presence metric—even when multiple such pronouns appeared within the same paragraph—thereby preserving proportional fidelity to narrative emphasis rather than syntactic density. The possessive pronoun 'mine' was intentionally omitted from the initial operational definition due to its relatively low frequency and higher potential for ambiguous referential scope in complex syntactic constructions.

Subsequent diagnostic review identified two methodological limitations requiring refinement: first, the original procedure did not fully account for extended direct speech spanning multiple consecutive paragraphs, potentially misattributing narratorial agency to quoted discourse; second, the exclusion of 'mine' introduced minor asymmetry in first-person lexical coverage. A revised protocol was therefore implemented to explicitly detect and exclude multi-paragraph quotation blocks using typographic and syntactic boundary heuristics, and 'mine' was formally integrated into the pronoun inventory. Empirical validation confirmed that neither adjustment altered the relative ranking order of candidates or affected the statistical significance of model outcomes, affirming the robustness of the core methodology. Notably, narrator-pronoun resolution was not applied to *The Moonstone*, given its distinctive epistolary structure involving six distinct narrators with divergent perspectives and temporal scopes, nor to the two Van Dine novels, both of which employ detached observer narrators whose discursive positioning precludes direct involvement in the investigative process.

Three distinct ranking paradigms were comparatively evaluated to assess their efficacy in identifying the central narrative agent. Method M0 relies exclusively on degree centrality—a topological measure quantifying the number of unique character

connections per node—assigning higher scores to figures embedded in denser relational networks. Method M2 adopts a lexical frequency baseline, ranking characters solely by raw mention count, thus serving as a simple yet informative benchmark for surface-level salience. Method M1 constitutes a supervised, multivariate model trained on the development subset, integrating seven complementary features: chronological position of first appearance, normalized mention frequency, weighted connection strength, degree centrality, betweenness centrality, eigenvector centrality, and detective tie—a normalized ratio capturing the proportion of a character's co-occurrences that involve the principal investigator. All features were transformed into within-novel percentile ranks to ensure scale invariance across heterogeneous textual lengths and densities. An aggregate culprit profile was then derived by computing the mean percentile vector across all culprits in the development set. Test-set candidates were ranked by ascending Euclidean distance from this reference profile. Betweenness centrality calculations employed binary edges restricted to pairs co-occurring in at least two paragraphs, whereas eigenvector centrality utilized weighted edges reflecting cumulative co-occurrence counts.

Model performance was assessed using three complementary metrics: top-1 accuracy (exact match of the highest-ranked candidate), top-3 accuracy (inclusion of the correct agent within the top three positions), and mean reciprocal rank (MRR), which provides a graded measure of ranking quality by averaging the inverse of each correct agent's position. In cases where a novel featured dual central agents—arising from collaborative or structurally parallel resolutions—the evaluation prioritized the agent receiving the higher numerical rank. Identical adjudication logic governed both random baseline simulations and the 20,000 one-sided permutation tests conducted to establish statistical confidence thresholds. Method M1 underwent dual validation: primary evaluation on the held-out test set and secondary assessment via leave-one-out cross-validation across the development novels to verify internal consistency. In contrast, M0 and M2—being unsupervised and parameter-free—were computed uniformly across the full corpus of twenty novels without partitioning, enabling direct comparison of baseline performance against the trained model under identical distributional conditions.

4. Results

The trained model designated as M1 demonstrated limited effectiveness in identifying the culprit across the evaluation set [6, 8]. When tested on the twelve held-out novels, it placed the correct perpetrator in the top-ranked position only once, yielding a top-1 accuracy of 8%. Its top-3 accuracy stood at 17%, and its mean reciprocal rank (MRR) was calculated at 0.25. For comparative context, a baseline random ranking strategy would yield theoretical expectations of approximately 12% for top-1 accuracy, 33% for top-3 accuracy, and an MRR of 0.31—indicating that M1 underperformed even relative to chance on the most stringent metric. In contrast, the simpler baseline model M0 achieved markedly stronger performance, attaining 25% top-1 accuracy and 58% top-3 accuracy. The mention-based model M2 further improved upon this, reaching 25% top-1 accuracy and 75% top-3 accuracy—suggesting that raw frequency of character mention carries substantial discriminative power in this task. Notably, M1's top-3 accuracy rose to 35% and its MRR to 0.349 when evaluated under a refined configuration; however, statistical testing revealed no significant difference from chance for the MRR metric ($p = 0.201$), whereas M0 and M2 achieved 55% and 65% top-3 accuracy respectively—highlighting the persistent gap between learned structural representations and surface-level lexical cues in this domain.

Hypothesis H2 investigates the distribution of culprits across positions of narrative prominence, specifically whether perpetrators tend to occupy highly visible roles among characters introduced early or frequently in the text. Analysis across thirteen novels revealed that the culprit appeared among the three most frequently mentioned characters in 65% of cases—a figure substantially exceeding the 31.4% expected under uniform random distribution ($p = 0.002$). Further decomposition showed that the second and third positions in the mention-ranking hierarchy were especially overrepresented: ten culprits

occupied either of these two slots, compared with only approximately 4.1 occurrences expected by chance ($p = 0.003$). Conversely, only three culprits ranked first by mention count, slightly above the 2.2 expected under randomness [6, 9]. This pattern suggests neither systematic avoidance nor preferential placement at the absolute apex of visibility, but rather a consistent tendency toward high—but not maximal—narrative salience. The phrase 'prominence without primacy' captures this nuanced distributional behavior, emphasizing that elevated attention is reliably present while strict dominance in mention frequency is not required or typical. It is important to clarify that the second component of this phrase—'without primacy'—is descriptive rather than inferential; it reflects an observed absence of first-place concentration rather than a statistically confirmed effect of deliberate demotion or suppression (As shown in Table 1).

Table 1. Culprit-identification Performance. Significance Values Are Specific to the Indicated Metric.

Model	Forward top-1	Forward top-3	Forward MRR	20-novel top-1	20-novel top-3	20-novel MRR
Random	12%	33%	.31	11%	31%	.30
M0	25%	58%	.46	20%	55%	.43
M1	8%	17%	.25	15%	35%	.35
M2	25%	75%	.50	15%	65%	.41

Temporal positioning of the culprit within the narrative structure also reveals distinctive patterns. Thirteen of the twenty-four culprits made their first appearance within the initial 10% of the novel's textual length, and seventeen appeared before the conclusion of the first quarter [10, 11]. The median point of first occurrence across all culprits was found at 6.0% of the total text, whereas innocent candidates appeared slightly later, with a median first appearance at 8.5%. However, this apparent temporal advantage diminishes when normalization is applied within each individual novel: when characters are ranked solely by order of introduction within their respective texts, only 20% of culprits ranked among the three earliest-introduced characters—falling below the 31% expectation under randomness ($p = 0.918$). This indicates that while early introduction is indeed a widespread convention in detective fiction—consistent with longstanding compositional norms emphasizing exposition and character establishment—it does not function as a reliable differentiator between guilty and innocent parties. In other words, early entry into the narrative is a shared feature of major characters broadly, rather than a diagnostic signal specific to the perpetrator. Consequently, readers cannot leverage chronological precedence alone to narrow the field of suspects meaningfully.

A more robust and statistically significant pattern emerges when examining character dynamics in the concluding segment of the narrative, as formalized in hypothesis H3. To quantify this, we computed a denouement ratio for each candidate character: the quotient obtained by dividing the proportion of mentions occurring in the final 20% of the text by the proportion occurring in the initial 80%. Across thirteen novels, the culprit ranked among the top three characters with the highest denouement ratio—a rate of 65% that significantly exceeds chance ($p = 0.002$); moreover, the mean reciprocal rank for this metric reached 0.555, with a highly significant p-value less than 0.001. This finding points to a consistent late-stage intensification of attention directed toward the perpetrator, suggesting that narrative resolution involves not merely revelation but also a retrospective amplification of the culprit's textual presence. Such a mechanism may serve both rhetorical and cognitive functions—reinforcing the solution through heightened salience while simultaneously enabling plausible misdirection earlier in the story by distributing attention more evenly across multiple characters prior to the climax.

To address potential methodological concerns regarding spurious inflation of the denouement ratio—particularly the possibility that characters with very low early mention counts could generate artificially large quotients due to minimal denominators—we conducted several robustness checks using alternative operationalizations. First, we verified that none of the twenty-four culprit entries had fewer than 23 explicit mentions

in the initial 80% segment, thereby excluding trivial or marginal cases. Second, substituting the ratio with an absolute difference between final-20% and initial-80% mention shares yielded comparable results: the culprit ranked first in ten novels and appeared among the top three in fourteen (70%), with $p < 0.001$. Third, imposing stricter minimum thresholds—requiring at least ten or twenty early mentions—still preserved the original 65% top-three result [2]. These convergent findings confirm that the observed effect is not attributable to statistical artifacts arising from sparse early data, but instead reflects a genuine and replicable shift in relative attention allocation during the narrative denouement. Nevertheless, it remains essential to acknowledge that this pattern constitutes a retrospective observation rather than a prospective cue: readers cannot compute such ratios in real time, as the ending has not yet unfolded, and thus this metric functions as an analytical lens for understanding authorial technique rather than a practical detection heuristic.

To facilitate deeper qualitative interpretation, we developed three exploratory typologies based on the magnitude and interpretive implications of the denouement ratio, jointly considering its numerical value and the character's established role in the preceding narrative. The 'under-exposure' category encompasses ten novels in which the culprit exhibited a denouement ratio of at least 2.0—indicating a doubling or greater increase in relative mention density during the final segment. This often corresponds to characters who were deliberately downplayed or backgrounded earlier, only to be foregrounded decisively at the resolution. The 'over-exposure' group includes six novels where the ratio did not exceed 1.10, suggesting minimal quantitative change in mention share—but where the ending nonetheless recontextualizes a character already prominent in the earlier narrative, transforming their perceived role through reinterpretation rather than increased visibility [3]. The remaining four novels fall into an 'intermediate or other' classification, reflecting more complex or ambiguous trajectories that resist binary categorization. These labels are intentionally interpretive and context-sensitive, grounded in close reading and narrative function rather than rigid algorithmic thresholds, and thus serve as conceptual scaffolds for literary analysis rather than empirically derived clusters.

A notable exception to standard network-based identification arises in the case of *The Murder of Roger Ackroyd*, where the narrator, Dr. Sheppard, occupies a structurally unique position. As the first-person narrator, he appears continuously throughout the text, and when personal pronouns such as 'I' and 'me' are systematically attributed to him, his centrality metrics become dominant: he ranks first in total mention count and network strength, and places at the 27th percentile for betweenness centrality—reflecting his pervasive connectivity across narrative segments. However, when analysis is restricted to explicit proper noun references only—excluding pronominal forms—the picture shifts dramatically: Dr. Sheppard drops to seventh place in mention frequency, eighth in strength, and remains at the 27th percentile for betweenness. This discrepancy underscores the critical importance of referential scope in computational literary analysis: models that ignore anaphoric reference risk severe misrepresentation of narrative agency, while those overly reliant on explicit naming may overlook constitutive elements of voice and perspective. The case illustrates how narrative mode—particularly first-person narration—introduces systematic asymmetries in textual representation that must be explicitly modeled to avoid distortion of character prominence assessments.

Sensitivity analyses further confirmed the stability of the core findings. Varying the reader-network construction threshold—specifically adjusting the minimum co-occurrence percentile cutoff from 70% to 80%—left M1's performance near chance levels, reinforcing the model's limited sensitivity to structural features under conservative parameter settings. At a more permissive 90% cutoff, M1's MRR improved to 0.40 ($p = 0.048$), though this gain coincided with increased inclusion of ending-dense passages, suggesting that the model benefits primarily from late-text information rather than from learning robust cross-sectional patterns [12, 13]. Excluding two outlier novels—*The Big Bow Mystery* and *Trent's Last Case*—produced consistent results: both the mention-frequency and denouement-ratio top-three rates settled at 61%, still significantly above

chance ($p = 0.007$). Additional perturbations—including modifications to narrator resolution protocols, adjustments to the co-occurrence window size (e.g., from five to fifteen tokens), and alternate handling of tied ranks in mention-based sorting—failed to alter the central interpretive conclusions. Collectively, these tests demonstrate that the observed patterns—especially the denouement-ratio effect and the prominence-without-primacy distribution—are resilient across methodological variations and therefore reflect stable properties of detective fiction composition rather than artifacts of specific analytical choices.

5. Discussion

A character may appear frequently across narrative passages, maintain extensive interpersonal connections within the story world, and even be tangentially associated with illicit activities—yet remain entirely uninvolved in the central crime. Social network analysis applied to literary texts captures patterns of co-occurrence and visibility, revealing which characters are most prominently situated in the textual fabric at various narrative stages; however, such structural metrics alone cannot determine culpability. The critical evidentiary elements required for reliable attribution of guilt or innocence—such as verifiable alibis supported by temporal consistency, corroborated witness testimony, access to confidential or privileged information known only to select individuals, and precise chronological sequencing of actions—are inherently absent from co-occurrence-based network representations [10]. Consequently, high centrality or elevated betweenness in a character network does not equate to criminal agency but rather reflects narrative salience, thematic function, or structural positioning within the plot architecture.

Once the narrative resolution unfolds and the true perpetrator is disclosed, analytical attention typically reorients toward that figure, often accompanied by retrospective reinterpretation of earlier textual cues [7]. This dynamic aligns with established theoretical frameworks concerning the dual-layered structure of detective fiction, wherein an initial surface narrative conceals a deeper, logically coherent sequence of events that only becomes fully intelligible upon revelation. While the concluding section inevitably foregrounds the culprit through explicit naming and sustained exposition, this heightened visibility is not merely a mechanical consequence of denouement—it functions as a deliberate rhetorical device that recalibrates readerly interpretation, inviting reassessment of prior interactions, dialogue subtext, and descriptive emphasis. Such narrative recentering underscores the genre's epistemological tension between appearance and reality, perception and evidence, and underscores how textual visibility operates independently of moral or legal accountability.

Narrative strategies for managing the identity of the perpetrator vary significantly across works: in some cases, the culprit remains deliberately obscured until the final chapters, exerting minimal discursive presence throughout the majority of the text; in others, the guilty party is introduced early and consistently present, yet their role is masked through narrative misdirection, unreliable narration, or strategic omission of motive and opportunity. Crucially, the conclusion does not merely identify a character—it transforms the reader's interpretive relationship with that figure, often recasting previously benign or ambiguous behaviors as symptomatic of concealed intent. The case of first-person narration further complicates network modeling: pronoun usage—particularly the distribution and clustering of 'I', 'me', and 'my'—introduces methodological challenges, as grammatical subjectivity does not map uniformly onto social agency or interactional frequency. Thus, quantitative network metrics must be interpreted alongside narratological context, discourse markers, and point-of-view constraints to avoid conflating linguistic prominence with causal responsibility.

Limitations and Future Research: Our investigation draws upon a constrained corpus of twenty English-language detective novels, limiting generalizability across broader literary traditions, historical periods, or authorial stylistic tendencies. While this sample permits preliminary identification of recurring structural patterns, it precludes robust comparative analysis of national publishing conventions, evolving genre conventions

over time, or author-specific narrative architectures. Furthermore, the adoption of the first 80% of each novel as a uniform temporal benchmark—though pragmatically useful for cross-textual alignment—fails to accommodate substantial variation in the timing of key revelations: some narratives disclose the perpetrator midway, others withhold identification until the final paragraph, and still others embed clues so densely that multiple plausible resolutions emerge. Additionally, the operational definition of character interaction based solely on paragraph-level co-occurrence overlooks crucial qualitative distinctions: shared physical space does not guarantee dialogue, observation, or influence; adjacency in prose may reflect editorial convention rather than narrative causality; and implicit references (e.g., off-stage mentions, hearsay, or hypothetical scenarios) introduce noise into interactional inference.

To strengthen methodological rigor and theoretical grounding, several expansions are warranted [14]. First, the corpus should be systematically diversified—not only in size but also in linguistic origin, publication era, and cultural context—to enable more nuanced exploration of how narrative networks reflect or diverge from real-world social dynamics. Second, inter-annotator reliability protocols must be implemented, involving multiple trained readers who independently code character interactions, motivations, and evidentiary weight, thereby mitigating subjective bias in network construction. Third, empirical validation through reader-response studies would significantly enhance ecological validity: controlled experiments could track how participants update suspect rankings at predetermined narrative milestones, measuring shifts in perceived likelihood of guilt, confidence levels, and reasoning strategies. Fourth, integrating multimodal features—such as speech act annotation, spatial mapping of scenes, and syntactic dependency parsing—would allow for richer, more granular modeling of character relationships beyond simple co-occurrence. Finally, developing adaptive temporal segmentation methods—calibrated to internal narrative markers like chapter breaks, investigative turning points, or forensic disclosures—would yield more accurate periodization for longitudinal network analysis.

6. Conclusion

This study systematically investigated the capacity of character network analytics to support narrative inference in detective fiction, with a specific focus on identifying the culprit prior to the narrative resolution. A comprehensive set of network-derived metrics—including centrality, betweenness, eigenvector scores, and co-occurrence frequency—was computed across multiple structural layers of the text, including dialogue adjacency, spatial proximity, and thematic association. Despite rigorous construction and iterative refinement of the character profile, the model failed to consistently prioritize the actual perpetrator above other plausible suspects; instead, simpler, more surface-level indicators—such as raw mention count, chapter-level appearance density, and syntactic prominence in exposition—proved more robust predictors of suspect ranking. Notably, visibility dynamics exhibited marked nonlinearity: characters who were structurally peripheral before the revelation often surged in network centrality post-reveal, while those initially dominant frequently experienced sharp declines in connectivity measures. This temporal shift reflects a deliberate narrative architecture wherein attentional scaffolding is reconfigured to align with epistemic reorientation. The concepts of under-exposure and over-exposure—defined here as systematic deviations from expected narrative salience given functional role—emerged as empirically observable patterns tied to concealment strategies, misdirection, and delayed exposition. While character networks do not function as deductive engines capable of solving the crime autonomously, they serve as sensitive diagnostic instruments for mapping the distribution, modulation, and recalibration of readerly attention across the narrative arc. Future work may extend this framework to cross-genre comparisons, multimodal integration with discourse markers, or fine-grained temporal segmentation aligned with plot beats.

Appendix. Per-Novel Data

This appendix presents corpus-level descriptive and inferential statistics essential for interpreting the primary analytical findings. The table summarizes aggregate outcomes across all novels included in the study, including distributional properties of culprit identification, ranking consistency, and denouement alignment. Rankings reflect comparative positioning between two culprits within dual-culprit narratives—specifically, the relative prominence assigned to the better-ranked culprit based on textual evidence such as frequency of attribution, narrative centrality, and resolution salience. Importantly, the exploratory classification framework extends beyond these two quantitative anchors; it incorporates qualitative dimensions—including syntactic framing, modality of accusation, temporal proximity to revelation, and functional role in plot closure—to capture nuanced patterns not reducible to rank order or ratio thresholds. These additional criteria were iteratively refined through close reading and intercoder calibration to ensure conceptual coherence and empirical discriminability across heterogeneous narrative structures. The resulting categories support hypothesis generation regarding how structural and rhetorical features jointly shape reader inference and interpretive consensus in detective fiction.

Table A1. Corpus, Culprit Ranking and Exploratory Classification.

Novel	k	c	Primary culprit	Mention rank	First rank	Ratio	Type
The Moonstone	12	1	Godfrey	8	6	1.026	Over
The Leavenworth Case	12	1	Harwell	4	8	0.467	Other
Mystery of a Hansom Cab	12	1	Moreland	4	1	3.559	Under
A Study in Scarlet	9	1	Hope	6	2	2.060	Under
The Sign of the Four	11	1	Small	5	3	3.074	Under
The Big Bow Mystery	8	1	Grodman	3	2	1.750	Other
The Hound of the Baskervilles	12	1	Stapleton	6	1	3.879	Under
Trent's Last Case	9	1	Cupples	6	2	1.575	Other
The Mysterious Affair at Styles	12	2	Alfred Inglethorp	4	6	1.055	Over
The Cask	12	1	Boirac	11	1	3.124	Under
The Red House Mystery	10	1	Cayley	4	6	0.444	Over
The Murder on the Links	12	1	Marthe	11	1	2.775	Under
Whose Body?	12	1	Freke	4	1	2.738	Under
The Benson Murder Case	10	1	Anthony Benson	2	2	3.615	Under
The Murder of Roger Ackroyd	12	1	Sheppard	1	5	0.969	Over
Unnatural Death	12	1	Mary Whittaker	4	4	1.380	Other
The Canary Murder Case	10	1	Spotswoode	4	1	7.648	Under
The Mystery of the Blue Train	12	2	Knighton	4	1	2.460	Under
The Seven Dials Mystery	12	2	Jimmy	1	2	0.908	Over
The Murder at the Vicarage	12	2	Lawrence Redding	11	5	0.867	Over

Table A2. Full-corpus Tests from 20,000 One-Sided Permutations.

Test	N	Top-1	Top-3	MRR	Random top-3	p(top-3)	p(MRR)
M0 most central (LOO)	20	20%	55%	0.427	31%	.023	.026
M1 culprit profile (LOO)	20	15%	35%	0.349	31%	.445	.201
M2 most mentioned / H2	20	15%	65%	0.412	31%	.002	.042
H1 first appearance	20	10%	20%	0.297	31%	.918	.492
H3 denouement ratio	20	35%	65%	0.555	31%	.002	< .001
H3 share difference	20	50%	70%	0.628	31%	< .001	< .001
H2 salience, excl. Big Bow & Trent	18	17%	61%	0.412	31%	.007	.043
H3 ratio, excl. Big Bow & Trent	18	39%	61%	0.562	31%	.007	< .001

Comprehensive candidate-level results—including statistical summaries for all 223 individual characters, their associated role designations (e.g., suspect, red herring, false accuser), and complete alias mappings—are provided in the supplementary materials accompanying this publication. These materials include annotated character matrices, cross-novel co-occurrence networks, temporal alignment heatmaps, and hierarchical clustering dendrograms derived from behavioral and linguistic feature vectors. Due to space constraints and the highly granular nature of these outputs, they are not replicated in the main text. Researchers may access the full dataset, codebook, and reproducible analysis scripts via the designated data repository, which adheres to FAIR principles (Findable, Accessible, Interoperable, Reusable) and includes version-controlled documentation for methodological transparency and longitudinal replication.

References

1. J. Seixas and A. Albet, "Introduction 1," in *Urban Governance in Southern Europe*, Routledge, 2016, pp. 1–11.
2. S. S. Van Dine, *Twenty Rules for Writing Detective Stories*. Booklassic, 2015.
3. F. Moretti, "Network theory, plot analysis," 2011.
4. D. Elson, N. Dames, and K. McKeown, "Extracting social networks from literary fiction," in *Proc. 48th Annu. Meet. Assoc. Comput. Linguistics*, July 2010, pp. 138–147.
5. V. Labatut and X. Bost, "Extraction and analysis of fictional character networks: A survey," *ACM Comput. Surv.*, vol. 52, no. 5, pp. 1–40, 2019.
6. T. Todorov, "The typology of detective fiction (1966)," in *Crime and Media*, Routledge, 2019, pp. 291–301.
7. M. P. Gelfer and J. F. Pazera, "Maximum duration of sustained /s/ and /z/ and the s/z ratio with controlled intensity," *J. Voice*, vol. 20, no. 3, pp. 369–379, 2006.
8. L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, pp. 35–41, 1977.
9. D. Altshuler, *Events, States and Times: An Essay on Narrative Discourse in English*. De Gruyter Open, 2016.
10. L. Frermann, S. B. Cohen, and M. Lapata, "Whodunnit? crime drama as a case for natural language understanding," *Trans. Assoc. Comput. Linguistics*, vol. 6, pp. 1–15, 2018.
11. V. Pojoga, "The Character Network in Detective Fiction: A Dialogic Analysis of Mihail Sadoveanu's The Hatchet," *Revista Transilvania*, no. 12, p. 5, 2025.
12. J. Barré, "Birth of the detective, birth of a genre," in *DH2026*, July 2026.
13. T. Kurwinkel, "The Commercial Genre Supersystem of Early Cinema Screen Detectives: On the Media Networks Surrounding Sherlock Holmes, Nick Carter, Stuart Webbs and Joe Deebs," in *German-Language Children's and Youth Literature In The Media Network 1900–1945*, Stuttgart: JB Metzler, 2023, pp. 481–497.
14. E. Bruno, L. Sabatino, and F. Tomasi, "FiCT-O: Modelling Fictional Characters in Detective Fiction from the 19th to the 20th Century," *Humanities*, vol. 14, no. 9, p. 180, 2025.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.