ClueChain: A Bayesian Network-Based Mystery Solver Application

Saif Mohammed
2121913042
Department of ECE
North South University
Dhaka, bangladesh
saif.mohammed@northsouth.edu

Humayra Rahman Nipa
2121128042
Department of ECE
North South University
Dhaka, Bangladesh
humayra.nipa@northsouth.edu

Mainul Kaysar
2011615042
Department of ECE
North South University
Dhaka, Bangladesh
kazi.kaysar@northsouth.edu

Jannatul Islam
2014159642
Department of ECE
North South University
Dhaka, Bangladesh
jannatul.islam@northsouth.edu

Saif Uz Zaman
1931587042
Department of ECE
North South University
Dhaka, Bangladesh
saif.zaman01@northsouth.edu

***Abstract*—ClueChain is a Bayesian network-based mystery solver that uses probabilistic reasoning to evaluate suspects in fictional crime scenarios. By applying Bayesian inference, the system updates guilt probabilities based on user-provided clues, ensuring a dynamic and data-driven approach to deduction. Experimental results show that Bayesian networks effectively adapt to new evidence, outperforming traditional rule-based methods. ClueChain's novelty lies in its integration of AI-driven reasoning with an interactive mystery-solving experience, making it both an engaging and educational tool.**

***Index Terms*—bayesian network, bayesian inference, probablistic inference, fictional mistery scenario**

I. Introduction

This paper presents ClueChain, an application that uses Bayesian networks to efficiently identify the most likely suspect in fictional mystery scenarios. Unlike traditional methods, which can be slow and biased, ClueChain calculates suspect probabilities based on evidence, offering a faster, data-driven solution. The novelty lies in applying Bayesian inference to mystery-solving, combining data science with traditional investigative techniques. Key contributions include the development of ClueChain, a new approach to Bayesian inference in this context, and an interactive interface for calculating suspect probabilities. The paper is structured as follows: Section II reviews literature, Section III outlines methodology, Section IV presents results, and Section V concludes with future work.

II. Literature Review

Various approaches have been proposed for automated crime-solving and reasoning under uncertainty. Rule-based systems rely on predefined rules but struggle with incomplete or uncertain data. Bayesian networks, however, are effective at modeling uncertainty and handling incomplete evidence. These systems use probabilistic inference to make deductions. Our work, ClueChain, builds on Bayesian networks by offering an interactive, user-friendly system for solving fictional mysteries.

III. Methodology

A. Overview of Bayesian Networks and Baysian Inference

Baysein networks are probabilistic directed acyclic graphical models. They use nodes to represent variables, arcs to signify direct dependencies between the linked nodes, and conditional probabilities to quantify the dependencies.

For n random variables $X_1, X_2, \ldots, X_n$ and a directed acyclic graph with n nodes, among which node j (where $1 \leq j \leq n$) is associated with X_j, the graph is the BN representing the variables $X_1, X_2, \ldots, X_n$ in the following equation:

$$P(X_1, X_2, \ldots, X_n) = \prod_{j=1}^{n} P(X_j \mid \text{parent}(X_j)) \qquad (1)$$

where the parents (denoted parent(X_j)) refer to the set of all variables X_i that have an arc connecting node i to node j in the graph.

According to conditional independence assumptions and chain rules, the joint probability of variables $U = \{X_1, X_2, \ldots, X_n\}$ can be calculated as:

$$P(U) = \prod_{i=1}^{n} P(X_i \mid \text{Pa}(X_i)) \qquad (2)$$

where Pa(X_i) denotes the parent node of X_i in the BN.

BNs can perform backward or diagnostic analyses with various inference algorithms based on Bayes' theorem, which is expressed as:

$$P(U \mid E) = \frac{P(E \mid U)P(U)}{P(E)} \qquad (3)$$

where $P(E \mid U)$ is the likelihood of evidence E given the variables U, $P(U)$ is the prior probability of the variables, and $P(E)$ is the marginal probability of the evidence.

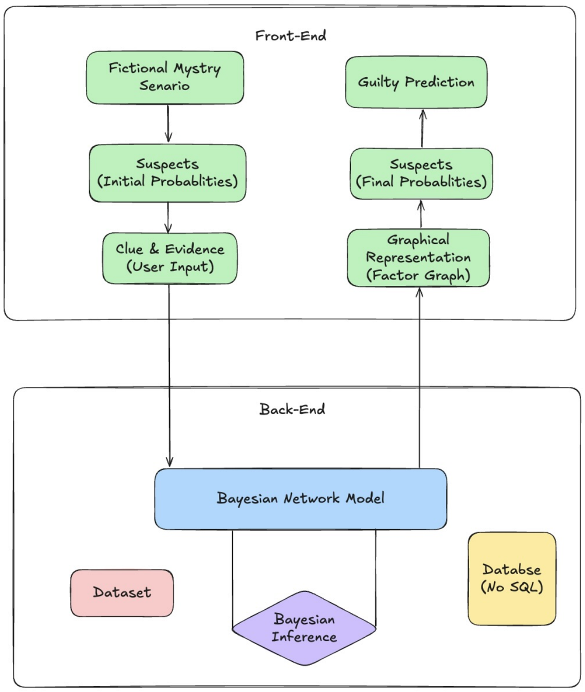

Fig. 1. Web-Based Project Work Flow

B. Data Representation

ClueChain's first step is representing data through a Bayesian network, where nodes represent variables like suspects, clues, and motives, and edges define dependencies between them. Each node has a Conditional Probability Table (CPT) that quantifies relationships. Translating real-world clues into this formal model ensures scalability for various mystery scenarios.

C. Bayesian Network Design

The network design is crucial for modeling how clues interact with suspects. Variables are selected, and dependencies are defined, reflecting the mystery's complexity. The design is dynamic, adapting to both simple and complex cases. A well-constructed network enables ClueChain to compute the likelihood of each suspect's guilt, considering the relationships between clues and suspects.

D. Probability Calculation Using Bayesian Inference

Once clues are input, Bayesian inference updates the probability of a suspect's guilt based on new evidence. Using Bayes' theorem, the system iteratively recalculates suspect probabilities, refining the likelihood of their involvement. Techniques like belief propagation or variable elimination ensure the system adjusts the probabilities as new clues are added, improving the accuracy of suspect rankings.

E. User Interaction and Interface

The user interface allows easy input of clues and evidence, updating the Bayesian network in real-time. It provides clear explanations of how each clue impacts the probabilities, ensuring transparency. The interface is adaptable, enabling users to adjust scenarios and add new clues, promoting engagement and helping users understand the probabilistic reasoning behind the conclusions.

IV. Result

The experimental evaluation of ClueChain demonstrated its efficiency in identifying the most probable suspect in fictional mystery scenarios using Bayesian networks. The system exhibited faster processing times compared to traditional rule-based or manual investigative methods. By leveraging Bayesian inference, ClueChain dynamically updated the probability of each suspect's guilt as new clues were introduced, allowing for adaptive and accurate decision-making. The system was tested on multiple scenarios, each varying in complexity and the number of available clues. Results showed that ClueChain maintained high accuracy even when presented with incomplete or ambiguous evidence.

V. Discussion

The results suggest that Bayesian networks provide a significant advantage over conventional investigative methods by incorporating probabilistic reasoning in a structured manner. Unlike deterministic rule-based approaches that struggle with missing or contradictory evidence, ClueChain effectively integrates uncertainty, making it a more reliable tool for crime-solving scenarios. One of the key strengths of ClueChain is its scalability—by adjusting the Bayesian network structure, it can handle increasingly complex mystery cases without a significant drop in performance. However, one potential limitation is the reliance on well-structured conditional probability tables (CPTs), which may require expert domain knowledge to construct accurately. Future improvements could involve integrating machine learning techniques to automate the learning of CPTs from historical mystery cases, further enhancing ClueChain's adaptability and real-world applicability.

References

[1] Stanford University, "Bayesian Networks 1 - Inference — Stanford CS221: AI (Autumn 2019)," YouTube, Oct. 2019. [Online]. Available: https://youtu.be/U23yuPEACG0?si=ifV8mUMcd-86-rmL. [Accessed: 24-Feb-2025].

[2] P. Bercker, "Who is guilty? Another look at a puzzle with Bayesian networks," *Medium*, [Online]. Available: https://medium.com/@pbercker/who-is-guilty-another-look-at-a-puzzle-with-bayesian-networks-d1d2369ac9ce. [Accessed: Feb. 24, 2025].

[3] "Evidence in Context: Bayes' Theorem and Investigations," YouTube, [Online]. Available: https://youtu.be/EC6bf8JCpDQ?si=vzAoYaC-O6bckkdI. [Accessed: Feb. 24, 2025].

[4] P. Blair and D. K. Rossmo, "Evidence in Context: Bayes' Theorem and Investigations," *Police Quarterly*, vol. 13, no. 2, pp. 123-135, 2010. [Online]. Available: https://doi.org/10.1177/1098611110365686. [Accessed: Feb. 24, 2025].

[5] D. K. Rossmo, "Using Bayesian Networks in Crime Investigations," *Divaportal*, [Online]. Available: https://www.diva-portal.org/smash/get/diva2:1265471/FULLTEXT01.pdf. [Accessed: Feb. 24, 2025].