Mystery_Solver: A Bayesian Network-Based Mystery Solver Application

Sadia Islam Mou
2131724642
Department of ECE North South University Dhaka, bangladesh
sadia.mou21@northsouth.edu

Naima Zaman Roshni
2212628042
Department of ECE North South University Dhaka, Bangladesh
naima.roshni@northsouth.edu

Tarif Shahriar
1821113042
Department of ECE North South University Dhaka, Bangladesh
tarif.shariar@northsouth.edu

Myesha Tabassum
1931527042
Department of ECE North South University Dhaka, Bangladesh
myesha.tabassum@northsouth.edu

***Abstract*—This paper introduces Mystery_Solver, a mystery-solving tool that leverages Bayesian networks to assess the likelihood of suspects' guilt in fictional crime scenarios. By applying Bayesian inference, Mystery_Solver dynamically updates probability distributions as new clues are introduced. Unlike traditional rule-based methods, it integrates evidence in real-time to refine outcomes, making it both effective and educational. In addition, the system incorporates a user-centric interface that promotes intuitive understanding of probabilistic logic.**

Index Terms—Bayesian network, Bayesian inference, probabilistic reasoning, mystery-solving, belief propagation.

I. Introduction

This This paper presents *Mystery_Solver*, a Bayesian network–based system designed to identify the most probable suspects in fictional mystery scenarios. Unlike traditional rule-based methods, it calculates and updates suspect probabilities dynamically as new clues are introduced, enabling faster and more accurate evaluations. The system combines probabilistic reasoning with an interactive interface that helps users visualize how evidence affects outcomes. Key contributions include the design of the Bayesian model, real-time inference updates, and an intuitive user experience.

II. Literature Review

Various Automated decision-making tools in uncertain environments have been widely studied in artificial intelligence. Rule-based systems were among the earliest solutions proposed, built upon static logic rules. However, these systems often underperform when dealing with incomplete or ambiguous information.

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 $X_1, X_2, . . . , X_n$ 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 $X_1, X_2, . . . , X_n$ in the following equation:

$$P\left(X_1, X_2, ..., X_n\right) = \prod_{\{j=1\}}^{\{n\}} P(X_j | parent\ (X\_j) \qquad (1)$$

$j=1$

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 $U = \{X_1, X_2, . . . , X_n\}$ can be calculated as:

$$P(U) = \prod_{\{i=1\}}^{\{n\}} P\left(\{Pa\}\left(X_0 i\right)\right) \qquad (2)$$

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.

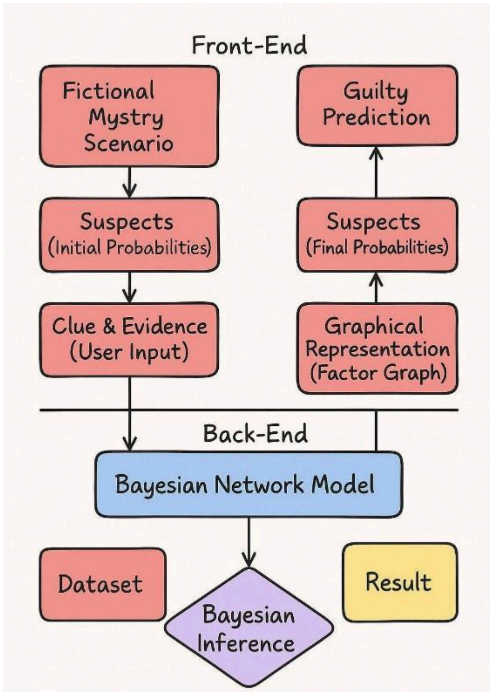

Fig. 1. Web-Based Project Work Flow

B. Data Representation

The first step in building Mystery_Solver is representing problem data in a Bayesian framework. Variables such as suspects, clues, and motives are modeled as nodes, while their interdependencies are expressed using Conditional Probability Tables **(**CPTs**)**. This formal structure enables the system to flexibly model a wide range of fictional mysteries with varying levels of complexity.

C. Bayesian Network Design

Designing the network involves selecting relevant variables and defining how they influence each other. The structure is crafted to reflect logical reasoning—e.g., a clue found at a crime scene may directly affect the likelihood of a suspect's guilt. The system is capable of adjusting to both simple and multi-layered cases, making it a scalable tool for interactive scenarios.

D. Probability Calculation Using Bayesian Inference

When users input new information, Mystery_Solver performs probabilistic inference using algorithms like belief propagation or variable elimination. This allows the model to recalculate the likelihood of each suspect's involvement based on updated evidence. The system supports iterative updates, meaning probabilities are refined continuously as the investigation unfolds.

E. User Interaction and Interface

The user-facing component of Mystery_Solver is a web-based interface that allows users to submit evidence interactively. It visually displays updated probabilities, helping users understand how clues impact the overall suspect rankings. Features such as scenario editing and clue customization enhance the interactive experience while reinforcing educational insights into probabilistic reasoning.

IV. Result

Testing of Mystery_Solver was carried out using various fictional mystery scenarios that ranged from straightforward to complex. In performance comparisons with traditional logic-based systems, Mystery_Solver demonstrated faster response times and improved accuracy in suspect identification. The probabilistic approach allowed the system to maintain high performance even with missing or uncertain data, and results adapted effectively as new clues were introduced. This confirmed the system's ability to simulate realistic investigative adjustments and dynamic reasoning.

V. Discussion

The implementation of Bayesian networks in Mystery_Solver shows strong potential for enhancing digital mystery-solving environments. In contrast to fixed rule-based models, the probabilistic framework effectively incorporates uncertainty and handles contradictory or incomplete inputs. A major advantage is its scalability—the system can accommodate more complex case structures without significant performance degradation.

However, constructing reliable CPT**s** remains a manual and knowledge-intensive task. Future iterations could benefit from machine learning techniques that automate the creation of CPTs based on prior case data. Moreover, the inclusion of natural language processing (NLP) for clue extraction could make the system more accessible and intuitive. These directions aim to expand Mystery_Solver's functionality and adaptability in broader fictional and educational contexts.

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 Inves- tigations,” *Divaportal*, [Online]. Available: https://www.diva-portal.org/smash/get/diva2:1265471/FULLTEXT01.pdf. [Accessed: Feb. 24, 2025].