

Abstract

In probabilistic inference, which predicts the likelihood of an event based on data and the relationships among events, a lack of data increases the uncertainty of the prediction, affecting decision-making depending on it. This study introduces a technique for **quantitatively computing the "variance of inference outcome,"** an indicator of this uncertainty. It has been difficult to compute due to the enormous computational burden. To address this, we use the technique of counting the number of solutions that satisfy given logical formulas. Our study can be used to **make more robust decisions that could otherwise lead to serious consequences.** For example, in predicting availability of infrastructure equipment, even if the prediction satisfies safety standards, a large variance value indicates high uncertainty. In such cases, our approach enables operational decisions such as encouraging additional data collection to avoid excess failure rate that exceeds the prediction.

Probabilistic Inference and Its Uncertainty

Probabilistic inference predicts the probability of occurrence of an event from data and the relationship among events.

- Blackout will increase the probability of disconnection.
- How much is the probability of disconnection considering both normal and blackout situations?

	Prob. of situations	Prob. of disconnection	
Normal	99%	0.5% when normal	Normal & disconnected $99\% \times 0.5\% = 0.495\%$
Blackout	1%	30% when blackout	Blackout & disconnected $1\% \times 30\% = 0.3\%$
			Disconnection = 0.795%

Predicted probability (**inference outcome**) will be used for, e.g., decision-making.

When our goal is $\leq 1\%$ prob. of disconnection,
We need not update equipment because the inference outcome is below the goal.

Lack of data causes uncertainty in the parameters used for inference. This leads to an **uncertain inference outcome.**

	Prob. of situations	Prob. of disconnection	
Normal	$(99 \pm 0.5)\%$	$(0.5 \pm 0.2)\%$ when normal	Disconnection $(0.795 \pm ??)\%$
Blackout	$(1 \pm 0.5)\%$	$(30 \pm 14)\%$ when blackout	

Since uncertainty in the inference outcome could affect the decision-making, we consider computing the **"variance of inference outcome,"** an index for this uncertainty.

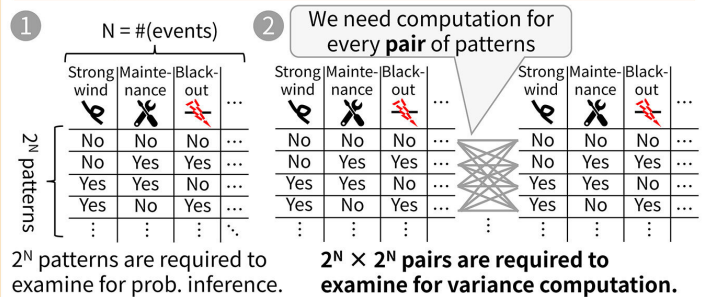
If the prob. of disconnection is $(0.795 \pm 0.1)\%$,
We need not update equipment because the inference outcome is safely below the goal.

If the prob. of disconnection is $(0.795 \pm 0.4)\%$,
We **should** update equipment because there's some possibility that the prob. of disconnection exceeds 1%...

Difficulty of Variance Computation

- 1 For probabilistic inference, we need to examine all patterns on the occurrences of the related events.
- 2 To compute the variance of inference outcome, we need more complicated calculation on these patterns.

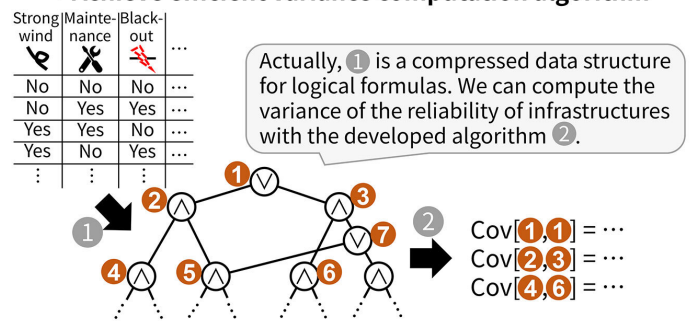
Traditionally, we can compute the variance in a reasonable time only when **the number of related events is up to 10.**



Solution by Counting on Logical Formulas

We proposed an algorithm to compute the variance in a reasonable time **even when #(related events) ≥ 100 .**

- 1 Transform the patterns on the occurrences of related events into a compressed data structure
 - 2 Develop an algorithm to compute the variance of inference outcome on the built data structure
- **Achieve efficient variance computation algorithm**



References

[1] K. Nakamura, M. Nishino, N. Yasuda, "Variance computation for weighted model counting with knowledge compilation approach" , in Proc. The 40th Annual AAAI Conference on Artificial Intelligence (AAAI ' 26), 2026.

Contact

Kengo Nakamura, Linguistic Intelligence Research Group, Innovative Communication Laboratory