Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)

Complex integrity constraint discovery: measuring trust in modern intelligent railroad systems

Wen-tao HU, Da-wei JIANG, Sai WU, Ke CHEN, Gang CHEN

Key words: Modern Intelligent Railroad Systems; Integrity Constraint; Model Reliability; Streaming data; Data consistency analysis

Cite this as: Wen-tao HU, Da-wei JIANG, Sai WU, Ke CHEN, Gang CHEN, 2022. Complex integrity constraint discovery: measuring trust in modern intelligent railroad systems. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, 23(10):832-837. https://doi.org/10.1631/jzus.A2200156

Introduction

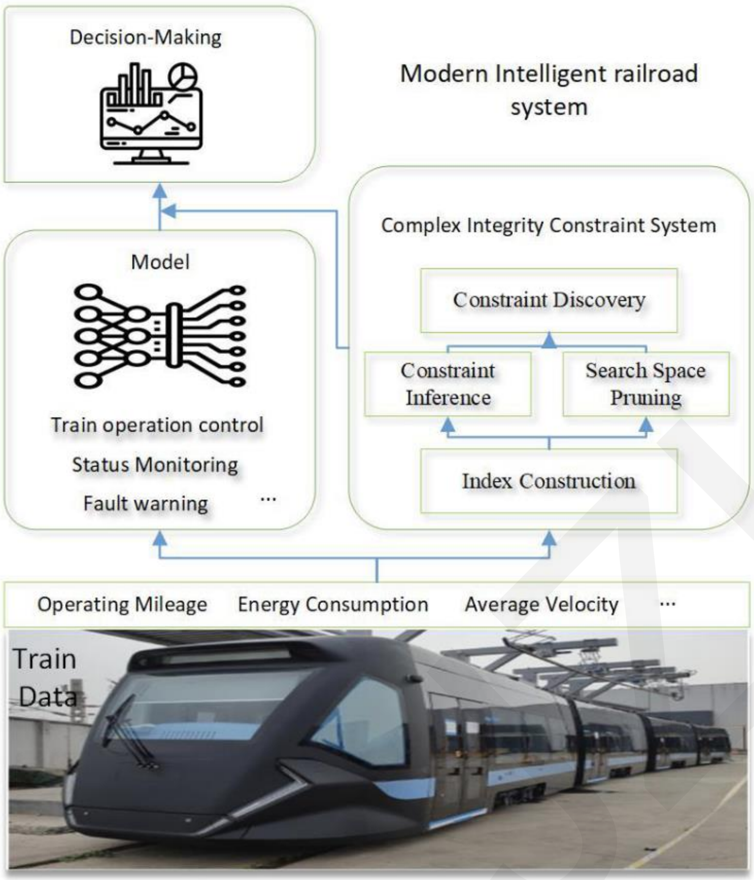
Motivations:

- The mechanism for assessing the trustworthiness of intelligent rail system inferences is of paramount importance.
- Train systems need a technical metric to ensure the reliability of **effectiveness** and **safety** of the model.

Technical Challenges:

- Lack of labeled data and metrics to analyze the credibility of model inference results.
- Technologies for high-speed train systems need to ensure real-time and security.
- Model plausibility analysis techniques require a certain level of interpretability in addition to handling multiple types of complex data such as categorical, numerical, and time-series data.

System Structure

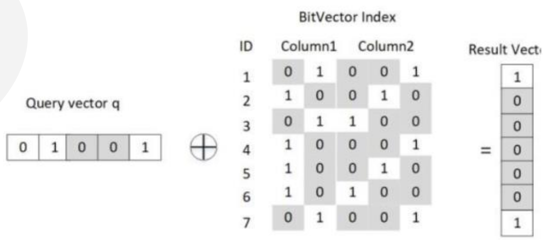


Index Construction

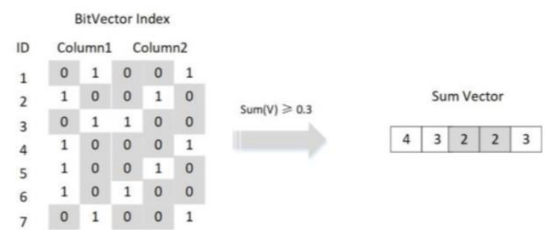
Purpose:

Through data vectorization, the system constructs **query** and **calculation** operators to accelerate the computation of support of attribute values as well as distribution characteristics.

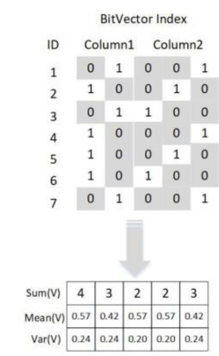
Query Example:



Calculation Example:

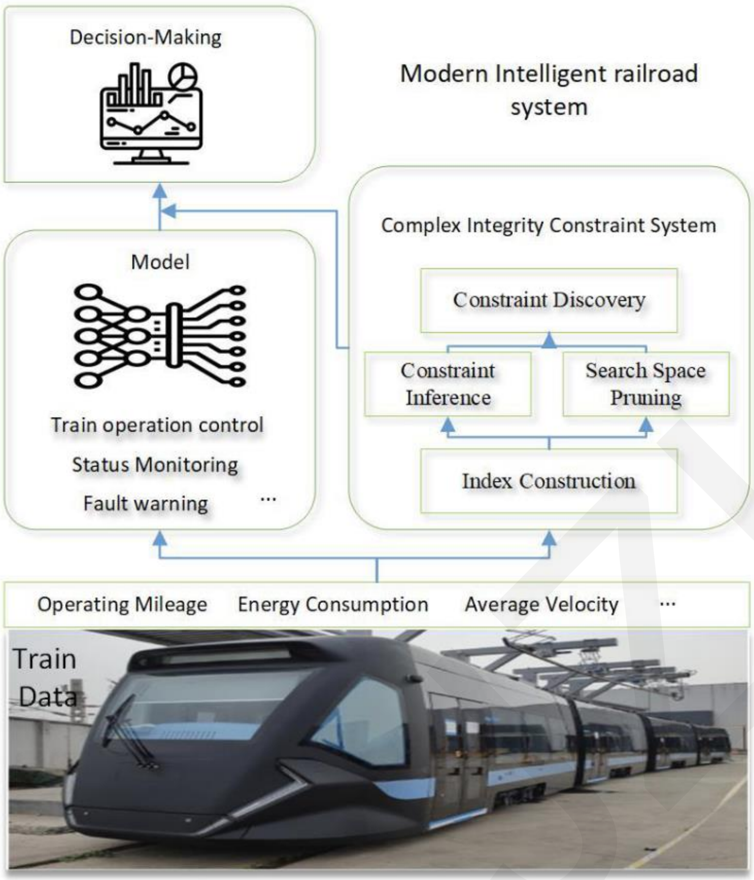


(a) Sum



(b) Mean and Variance

System Structure



Search Space Pruning

Purpose:

Through pruning techniques, candidate constraints that do not meet the threshold are **ignored** in advance, thus reducing the search space to improve the efficiency of constraint discovery.

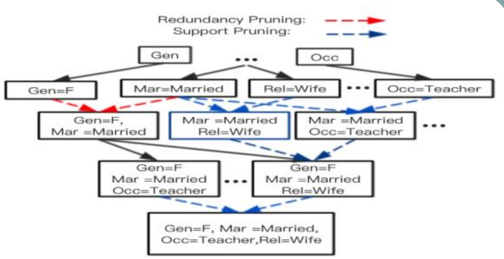
Formulation

$$Support(X) = \frac{|Dom(X)|}{n} \qquad Support(X_1, \dots, X_n) = \min(Support(X_1), \dots Support(X_n))$$

where $|Dom(X)|$ denotes the number of different attribute values for attribute X. n is the total number of tuples in the dataset.

Pruning Example:

The blue arrows indicate that the attribute conditions **do not satisfy the threshold**, so these candidate constraints **do not require** further computation.



Trusted Machine Learning

Unsafe tuple:

Given a class C of functions, and a labeled instance $[r;Y] \subset [Dom(r) \times Dom(Y)]$ a tuple $t \in Dom(r)$ is an unsafe tuple .
w.r.t. C and $[r;Y]$ if $\exists\ f,g \in C\ s.t. f(r)=g(r)=Y$ but $f(t) \neq g(t)$.

Proposition:

There exists a complex integrity constraint Φ for r s.t. The following statement is true: "∃Φ(t) $iff\ t$ is unsafe w.r.t C and $[r;Y]$ for all $t \in Dom(r)$".

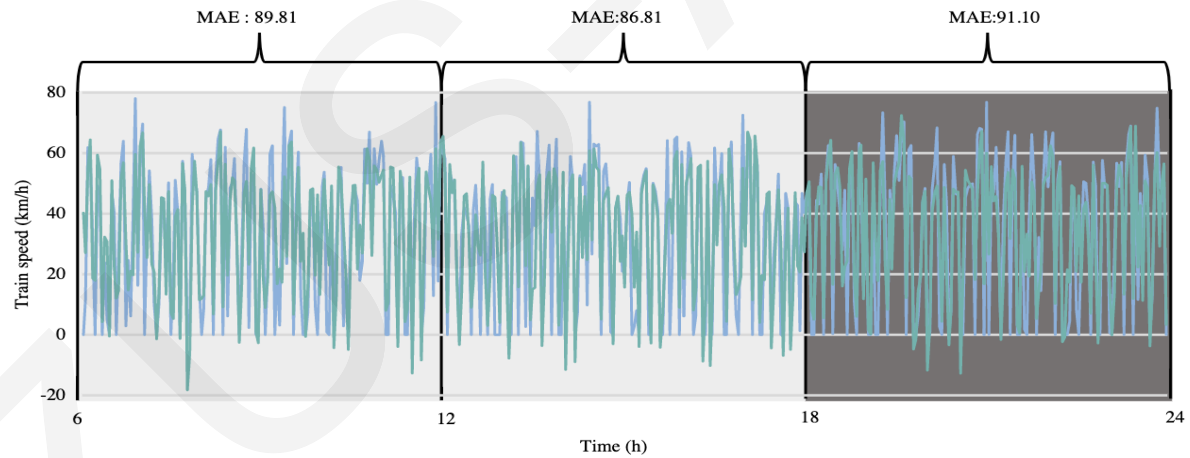
Performance Evaluation and Discussion

Evaluation Metrics:

Mean absolute error (MAE)

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

Results: Comparison of real and model-inferred speeds of trains at different time periods



Discussion:

- The results of the model inference from 12-18 hours are closer to the real data (minimum MAE) compared to other time periods, and the average percentage of violation data in this period is also the smallest.
- The violation of constraints is a very good feature of prediction errors.

Conclusions

- We present a system for automatically discovering complex integrity constraints.
- The concept of unsafe tuples for trusted machine learning.
- The system uses the BitVector index structure to speed up the computation of sums, means, and variances.
- Experiments validate our theory that complex integrity bounds provide an efficient and robust mechanism to quantify the trust of inferences from models deployed by intelligent railroad systems on data