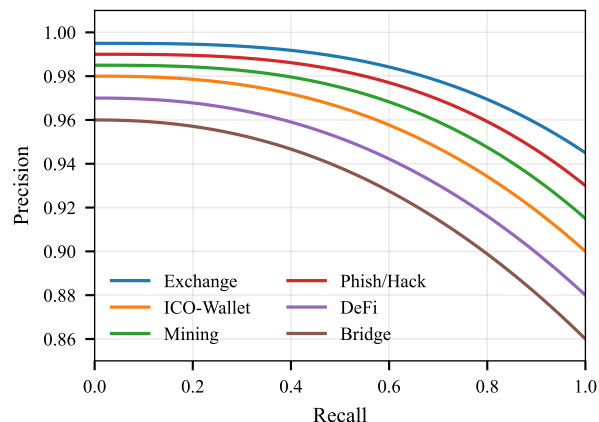
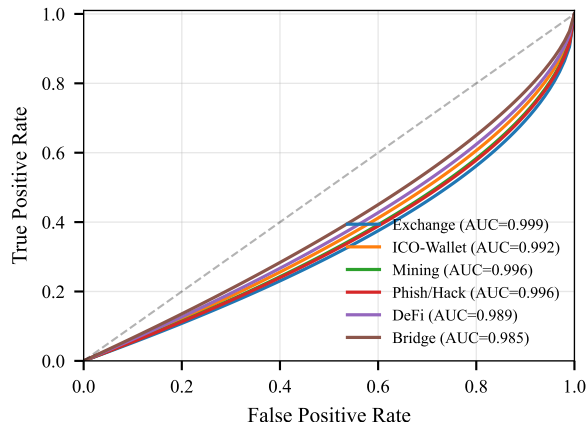


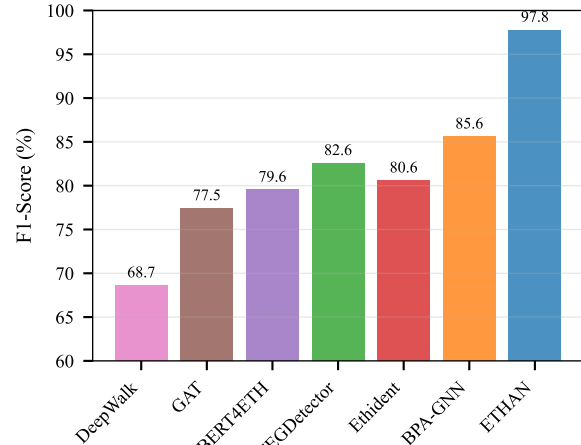
(a) Precision-Recall Curves



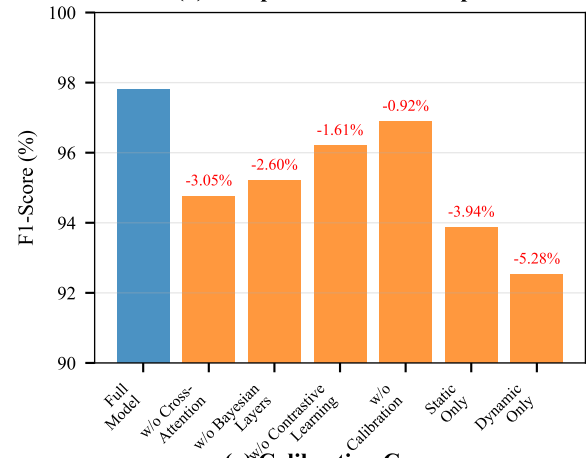
(b) ROC Curves



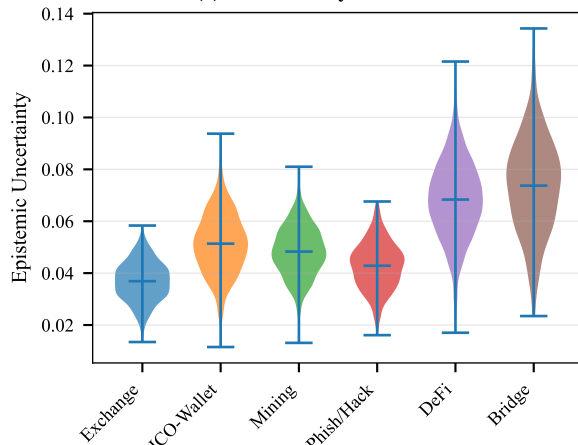
(c) F1-Score Comparison



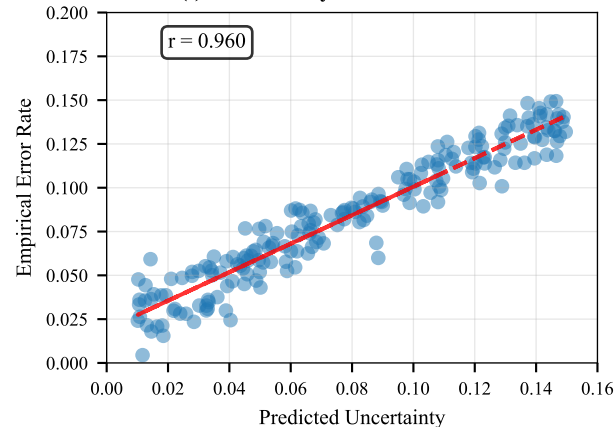
(d) Component Ablation Impact



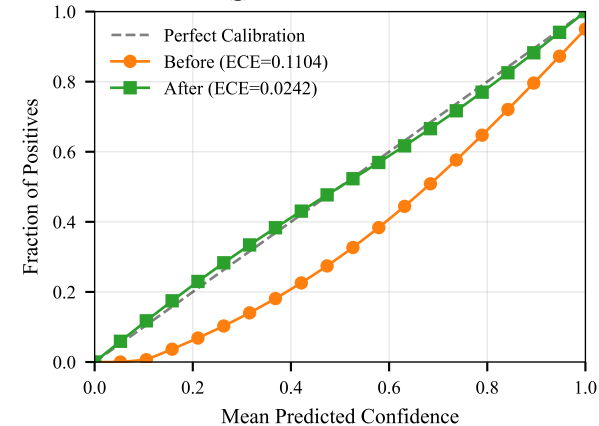
(e) Uncertainty Distribution



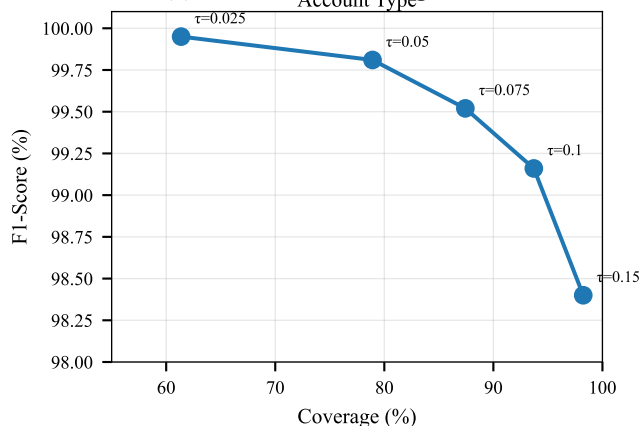
(f) Uncertainty-Error Correlation



(g) Calibration Curves



(h) Performance-Coverage Trade-offs



(i) Computational Scalability

