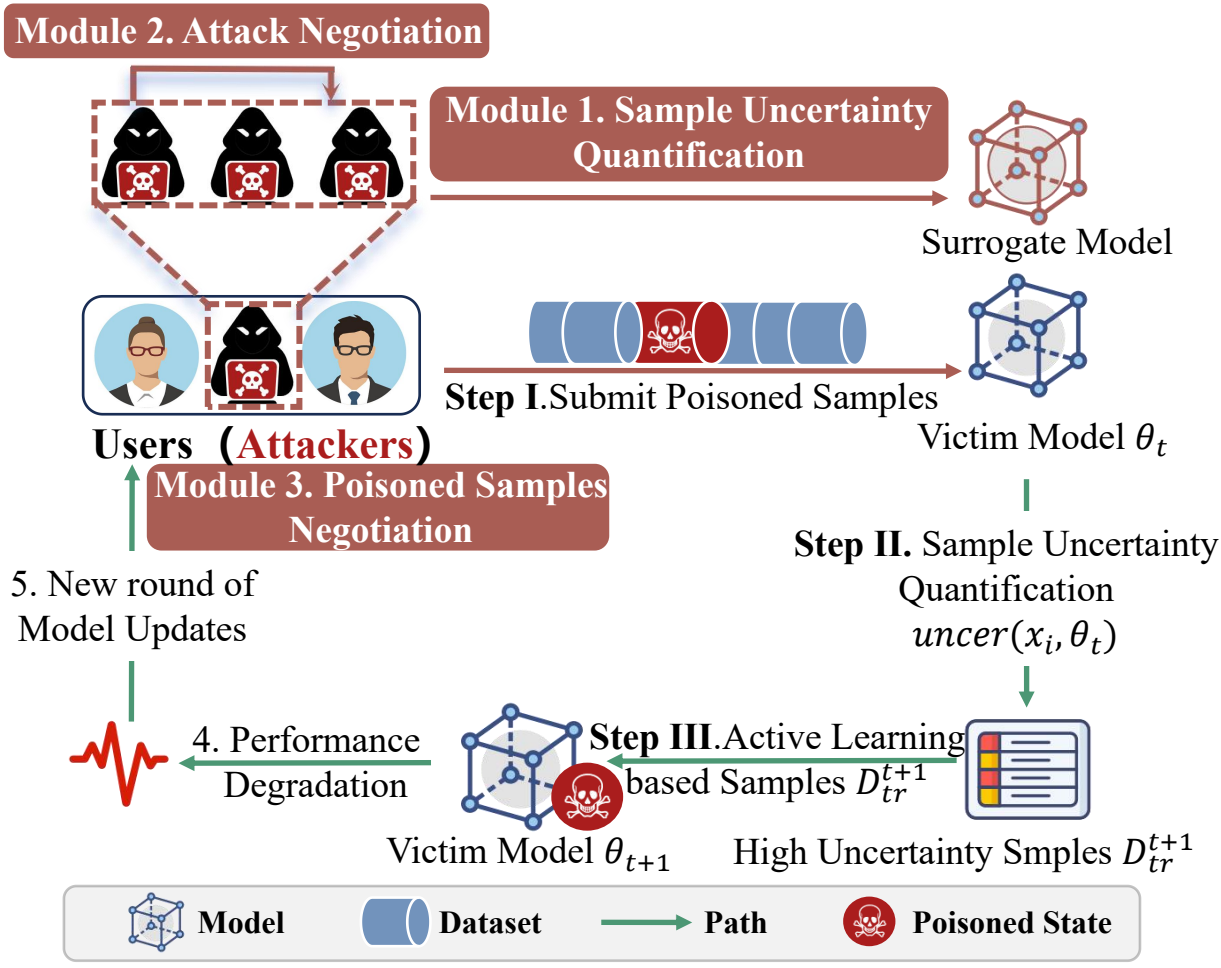
Module 2. Attack Negotiation
Module 1. Sample Uncertainty Quantification
Surrogate Model
Step I.Submit Poisoned Samples
Victim Model θ_t
Users (Attackers)
Module 3. Poisoned Samples Negotiation
5. New round of Model Updates
Step II. Sample Uncertainty Quantification
$uncer(x_i, \theta_t)$
4. Performance Degradation
Step III.Active Learning based Samples D_{tr}^{t+1}
Victim Model θ_{t+1}
High Uncertainty Smples D_{tr}^{t+1}
Model
Dataset
Path
Poisoned State