Motivation

• More and more innovative algorithmic computations are being deployed to the public cloud in important applications (e.g., business analytics, geospatial mapping/searching, bioinformatic analysis, and image processing).
• Control flow, which decides the sequence of instructions (e.g., statements) to be executed, directly reflects the algorithm of a program.
• How to protect control flow confidentiality (and thus the confidentiality of the algorithm) of an outsourced program deployed to the public cloud?

Solution: Runtime Control Flow Obfuscation

Idea:

• Transform each original program into a public program and a private program.
• Public program: execute on the public cloud; perform most execution except for evaluating predicates of branch statements.
• Private program: execute on private cloud; evaluate predicates.

Proposed Techniques:

• Replace predicates of the branch statements in the original program with control flow query (CFQ) invocations to the private program.
• Insert indistinguishable fake branch statements in the original program to raise the bar for the attacker to understand the algorithm’s control flow.
• Maintain a continuous cache to reduce the cost of cross-cloud communication.

Transform original predicates

Predicate Normalization

Normalized inequality: \( x < y \)

Arithmetic Manipulation

Select predicate secret as 23, reorganize the inequality:

\[ x = (x + y + 28) < 23 \]

Transform to CFQ function call in public program

\[ \text{CFQ}(x + 28, L1) \]

Condition Substitution

Invoke

Private program

\[ \text{if}(\text{CFQ}(\text{exp}, \text{call_site_id}); \text{switch}(	ext{call_site_id}; \text{case} "L1") \text{return} \text{exp} \text{val} = 23; \}

Insert fake branch statements

Variable Range Eval.: (x > 6), (45 > y)

Expression Aggregation: x + 45 > 6 + y

Expression Relaxation: x + 45 + 2 > 6 + y

Choose Predicate Secret (7): x - y + 48 > 7

Transform to CFQ predicate; Insert fake branch statement

Predicate Normalization

Arithmetic Manipulation

Condition Substitution

Secret

Finsh

Returns true

Returns false

Continuous Cache

• For each \( \text{CFQ}(\text{exp}, \text{Ln}) \) on the public cloud, maintain a \text{true}_\text{bound} and a \text{false}_\text{bound} and implement a public program version.
• For each invocation of \( \text{CFQ}(\text{exp}, \text{Ln}) \) on the public program, if \( \text{exp} > \text{true}_\text{bound} \), the public program version returns true.
If \( \text{exp} < \text{false}_\text{bound} \), the public version returns false.
Otherwise, invoke on the real CFQ function in the private program.

Experiments Result

Execution time of MapReduce jobs with different obfuscation degree d

• The experiments are performed on a hybrid cloud (private cloud and Amazon Elastic MapReduce).
• Obfuscation degree: The probability of inserting a fake branch statement before each original statement.

RCFO Example

The original function of the K-means clustering algorithm

The function after RCFO: the constant variable CONVERGENCEDELTAl is hidden in the private program.