An Attribute-Based Access Control Model for Secure Big Data Processing in Hadoop Ecosystem

Maanak Gupta, Farhan Patwa, and Ravi Sandhu

Institute for Cyber Security,
Center for Security and Privacy Enhanced Cloud Computing,
Department of Computer Science
University of Texas at San Antonio

3rd ACM Workshop on Attribute-Based Access Control (ABAC’18)
Tempe, Arizona, USA, March 19-21, 2018
Outline

➢ Introduction and Motivation
➢ Multi-layer Hadoop Authorization Framework
➢ Object Tagged - RBAC Model
➢ HeABAC Model
➢ Implementation Approach
➢ Use Case
Big Data and Big Challenges

➢ IDC 2025:
  ❖ global “Datasphere” – 163 zettabytes
  ❖ 10x than 2016

➢ Security:
  ➢ Privacy Concerns (eg: HIPPA)
  ➢ Fine granular access requirements

➢ Hadoop Ecosystem = Hadoop core + Open-Source Projects
➢ Hadoop Data Lake
Hadoop Ecosystem
Authorization Architecture

Policy Manager: Apache Ranger, Apache Sentry
Gateway: Apache Knox
Ecosystem Service (ES): Apache Hive, HDFS, Apache Storm, Apache Kafka, YARN
Multi-Layer Access Control

Data and Service Objects

Services
- HDFS NameNode
- YARN ResourceManager
- Apache Hive

Objects
- HDFS Files
- Hive Tables
- Kafka Topics

Cluster Resources and Applications
- YARN Queues
- Cluster Nodes
HeAC Model: Consolidated View

Hadoop Ecosystem Access Control Model

Users (U), Groups (G), Subjects (S)
Hadoop Services (HS)
Hadoop Service Operations ($O_{HS}$)
Objects (OB), Operations (OP)
Ecosystem Service (ES), Objects (OB)
Operations (OP), Tag

© Maanak Gupta
OT-RBAC Model

Object-Tagged RBAC
HeABAC Model

Hadoop Ecosystem Attribute-Based Access Control Model

User Attributes (UA),
Hadoop Services Attributes (HSA)
Objects Attributes (OA),
Ecosystem Service Attributes (ESA)
Group Based Attribute Inheritance

Roles: {student}  
Roles: {student, Staff, Grader, TA}  
Roles: {TA}  
Roles: {Grader}  
Roles: {Staff, Grader, TA}

Group Hierarchy  

Major Benefits: Easy Administration where multiple roles can be assigned to user with single administrative operation.
HeABAC Use-Case

companyName: {mccombs}
address: {spring-well drive}
dealership: {toyota}

effective attributes

group membership

certification: {ASE}
department: {diagnostic}
role: {technician}
companyName: {mccombs}
address: {spring-well drive}
dealership: {toyota}

effective attributes

Access Control Point

subject

createdBy: {admin2}
serviceType: {DataNode}
nodeLabel: {Table1}
direct attributes

Access Control Point

request communication

policy retrieval and evaluation

ECOSYSTEM SERVICE AND OBJECT

hive
car1
contains
tableType: {sensor-data}
car: {FVR1234}
serviceType: {HIVE}

createdOn: {11-11-2017}
createdBy: {admin1}

direct attributes

direct attributes

Cross Service Trust

datanode

NameNode

HADOOP SERVICES

request select operation

© Maanak Gupta

World Leading Research with Real World Impact!
1. Authorization\textsubscript{access}(s:S, es:ES) \equiv \text{diagnostic} \in \text{effective}\textsubscript{department}(s) \land\text{technician} \in \text{effective}\textsubscript{role}(s) \land \text{serviceType}(es) = \text{HIVE} \land \text{createdBy}(es) = \text{admin}1.

2. Authorization\textsubscript{select}(s:S, es:ES, ob:OB) \equiv \text{Authorization}_{\text{access}}(s:S, es:ES) = \text{True} \land \text{diagnostic} \in \text{effective}\textsubscript{department}(s) \land \text{effective}\textsubscript{role}(s) \in \text{readerType}(ob) \land \text{tableType}(ob) = \text{sensor-data} \land \text{car}(ob) = \text{FVR1234}.

3. Authorization\textsubscript{access}(s:S, hs:HS) \equiv \text{diagnostic} \in \text{effective}\textsubscript{department}(s) \land \text{technician} \in \text{effective}\textsubscript{role}(s) \land \text{serviceType}(hs) = \text{DataNode} \land \text{createdBy}(hs) = \text{admin}2
HeABAC Administrative Realms
ARBAC inspired GURA, GURA_G models are required.
Conclusion and Future Work

➢ Hadoop Authorization Layers
➢ Object-Tagged-RBAC Model
➢ Formalized Attributes based HeABAC Model

Some Future Goals:
➢ Introduce Data ingestion security
➢ Privacy concerns and finer grained approaches in multi-tenant Hadoop Lake