Adaptive Reordering & Clustering Based Framework for Efficient XACML Policy Evaluation

Said Marouf, Mohamed Shehab, Anna Squicciarini, and Smitha Sundareswaran

***Abstract*— The adoption of XACML as the standard for specifying access control policies for various applications, especially web services is vastly increasing. This calls for high performance XACML policy evaluation engines. A policy evaluation engine can easily become a bottleneck when enforcing XACML policies with a large number of rules. In this paper we propose an adaptive approach for XACML policy optimization. We apply a clustering technique to policy sets based on the K-means algorithm. In addition to clustering we find that, since a policy set has a variable number of policies and a policy has a variable number of rules, their ordering is important for efficient execution. By clustering policy sets and reordering policies and rules in a policy set and policies respectively, we formulated and solved the optimal policy execution problem. The proposed clustering technique categorizes policies and rules within a policy set and policy respectively in respect to target subjects. When a request is received, it is redirected to applicable policies and rules that correspond to its subjects; hence, avoiding unnecessary evaluations from occurring. We also propose a usage based framework that computes access request statistics to dynamically optimize the ordering access control to policies within a policy set and rules within a policy. Reordering is applied to categorized policies and rules from our proposed clustering technique. To evaluate the performance of our framework, we conducted extensive experiments on XACML policies. We evaluated separately the improvement due to categorization and to reordering techniques, in order to assess the policy sets targeted by our techniques. The experimental results show that our approach is orders of magnitude more efficient than standard Sun PDP.**

***Index Terms*— Policy Evaluation; Policy Categorization; XACML;**

I. Introduction

An access control policy is a set of rules that enables resource owners and administrators to control access and dissemination of their shared resources. For example, the Bloomberg financial service access control policy would only allow subscribed users to access the stock market data for market related studies. With the growing number of internet services and users, this directly implies an increase in the number of access requests and in turn an increase in the number of policy evaluations. When a user requests access to a certain resource, the access control module evaluates the policy rules to decide whether to allow or to deny access to the requested resource. With the continuously expanding number of resources and the increasing diversity and size of online systems, policies are becoming more complex and will involve a large number of rules. For example, social network sites typically host about 100 million users [3], [16], with over 25 million photos uploaded daily. Efficient policy evaluation techniques are required to ensure that policy evaluation introduces low latency without affecting the correctness of the evaluation process. Taking the widely adopted XACML (eXtensible Access Control Mark-up Language) [15] policy as an example, a policy set is composed of a set of policies, where each policy is divided into a set of rules. XACML not only provides a formalism to specify authorization policies, but it also includes information useful in making authorization decisions, as well as approaches to integrate constraints specified by multiple subjects, such as the policy combination algorithm. The policy combination algorithm along with other features unique to XACML, make it a very flexible and rich language.

A policy in XACML is evaluated by an XACML engine. The XACML engine is essentially composed by two main components, the Policy Enforcement Point (PEP) and the Policy Decision Point (PDP). The PEP is in charge of receiving an access request and translating it into an XACML request. The PEP then sends the XACML request to the PDP, which stores user specified access control policies written in the XACML policy language. The PDP checks the request with a set of XACML policies, and determines whether the XACML request should be permitted or denied. The evaluation process, in turn, has two main phases: first the policy to be used is selected, and second the rule among those appearing in the policy is used. The designers of current XACML engines, however, have not taken into account performance of the policy evaluation process. For example, Sun XACML PDP [19], which is the first and most widely used evaluation engine, performs brute force searching by comparing a request with all the rules in an XACML policy. Clearly, this approach does not efficiently support a large number of users' requests, who need prompt access to the data they are entitled to. To enable an XACML policy evaluation engine to process simultaneous requests of large quantities in real time, especially in face of a burst volume of requests, an efficient XACML policy evaluation engine is necessary. Our work aims at providing such a policy engine.

Achieving this goal is a challenging task however, due to the complexity of both the XACML policies, and of the evaluation process. Policies need to be reorganized according to the incoming access request type, in a possibly inexpensive and adaptive manner. Considering the example of Bloomberg's financial service, the PDP may face bursts of huge volumes of requests from the registered traders to access a particular stock's rate when there is a rush to either buy or sell a stock. However, when the trading is sluggish, the users may be more interested in reviewing their portfolio. Each of these requests are targeted at different objects and the policies need to be reorganized adaptively based on which type of request is being received. The reordering process itself needs to be lightweight and shuffle both the policies and the rules composing them. Additionally, in order to preserve the original

Mr. Marouf and Dr. Shehab are at The University of North Carolina at Charlotte, {smarouf, mshehab}@uncc.edu

Dr. Squicciarini and Ms. Sundareswaran are at The Pennsylvania State University, {acs20, sus263}@psu.edu

Digital Object Indentifier 10.1109/TSC.2010.28

intention of the policy writers, it is important that the reordering process does not affect the policy evaluation results, that is, the response to access requests must not change.

Starting from the Sun policy evaluation engine [19], in this paper we present the design and implementation of a simple yet effective framework that greatly improves the performance of XACML policies evaluation. Our design draws from the following two observations: (1) users who share common properties have the same request types, thus, the same subset of rules are evaluated, and (2) optimal rule ordering is subjective to the actual user requests.

We formulate the rule optimization problem for access policy requests, and show that our usage framework solves it. Precisely, our problem consists of finding which policy is applicable to an incoming request and also optimizing the ordering of the rules within the policy to match the request. In order to allow for this type of matching, we propose a technique that utilizes actual users' requests' characteristics. We categorize the users' access requests at two levels. Based on observation (1) we first categorize the request by subjects to see which policy would be applicable to it. Then, based on observation (2) we find a match between the request and the *execution vectors* for that policy. Execution vectors are the order in which the rules in a policy are applicable to a request. We build execution vectors by using different statistics to evaluate the cost of a rule and their frequency, and develop an approach to efficiently reorder policies and rules based on the specific properties of access requests.

We implemented our proposed framework as an extension of the Sun PDP engine. We choose this engine due to its popularity and to the fact that it is open source. We conducted extensive experiments on synthetic XACML policies of large sizes. We tested policies of different structures and sizes, and conducted experiments using different sets of access requests. The experimental results show that our framework is orders of magnitude more efficient than Sun's PDP, and the performance difference between our and Sun's PDP grows almost linearly with the number of rules in XACML policies. We tested the categorization and reordering techniques separately, and find interesting results on how our categorization technique by itself already outperforms the Sun implementation by orders of magnitude. The reordering provides a means for adaptability to user requests to further enhance the performance of the policy evaluation subject to different request trends. The required preprocessing time, necessary for categorization and reordering of policies and rules, is negligible as compared to the substantial improvement obtained.

The main contribution of this paper are thus summarized as follows:

- We formulate the optimal execution problem for access requests, and provide novel techniques to solve the problem, by taking into account both the policy reordering process and the rule process, while preserving the correctness of the reordered policies.
- We provide a policy categorization mechanism to enable the efficient policy processing.
- We prove through extensive tests that our approach outperforms the Sun's implementation by orders of magnitude.

The rest of the paper is organized as follows. In the next section we present some background information on XACML and access requests. In Section III we present our usage framework, present the optimal rule ordering problem, and provide an efficient algorithm to reorder rules and policies. In Section IV, we present our categorization based optimization. Our experimental results are shown in Section V, whereas related work is discussed in Section VI. We conclude the paper with conclusions and pointers for future research directions in Section VII.

II. Preliminaries

In this section we provide the logic formalism adopted throughout the paper to denote XACML policies and access requests. XACML policies are composed of five basic components, namely, *PolicySet*, *Policy*, *Target*, *Rule*, and *Policy and Rule Combining algorithm* for conflict resolution. The root of the XACML policy is the PolicySet element, which is defined as follows:

Definition 1: PolicySet is a tuple $PS = (id, t, P, PC)$, where:

- id is the PolicySet id.
- t is the PolicySet Target element, and takes values from the set {Applicable, NotApplicable, Indeterminate}.
- $P = \{p_1, \ldots, p_n\}$ is the set of policies.
- PC is the policy combining algorithm.

A *Policy* element is a set of rules and conditions that control access to protected resources which we refer to as objects. A policy contains a *target*, a set of *rules*, and a *rule combining* algorithm.

Definition 2: A policy is a tuple $P = (id, t, R, RC)$, where:

- id is the policy id.
- t is the policy target element, and takes values from the set {Applicable, NotApplicable, Indeterminate}.
- $R = \{r_1, \ldots, r_n\}$ is the set of rules.
- RC is the rule combining algorithm.

The Target element t specifies a set of predicates on the request attributes, which must be met in a PolicySet, Policy or Rule to apply to a given request. The attributes in the target element are categorized into *Subject, Resource* and *Action*. The attribute values in a request are compared with those included in the Target, if all the attributes match then the Target's PolicySet, Policy or Rule is said to be Applicable. If the request and the Target attributes do not match then the request is NotApplicable, and if the evaluation results in an error then the request is said to be Indeterminate. If a request satisfies the target of a policy, then the request is further checked against the rule set of the policy; otherwise, the policy is skipped without further examining its rules. The Target predicates can be quite complex, and can be constructed using functions and attributes. The rule combining algorithm RC respectively allows one to specify the approach to compute the decision result of a policy when the policy contains rules evaluating to conflicting effects. The policy combining algorithm PC follows the same logic but at the PolicySet level.

A *Rule* identifies a complete and atomic authorization constraint that can exist in isolation with respect to the policy in which it has been created. We define rules as follows.

Definition 3: A Rule is a tuple $r = (id, t, e, c)$, where:

- id is the rule id.
- t is the rule target element, and takes values from the set {Applicable, NotApplicable, Indeterminate}.
- e is the rule effect, where $e \in \{Permit, Deny\}$.
- c is a boolean condition against the request attributes.

The rule target element is similar to the policy target instead it indicates the requests applicable to the rule. The condition c

is a boolean function with respect to the request attributes. The rule's effect e, which can be Permit or Deny, is returned if the rule's condition c evaluates to true. The rule evaluation can also be Indeterminate in case of an error, or NotApplicable if the rule's target doesn't apply to the request's attributes. Access requests are typically matched against a policy set. A policy set is the root of an XACML policy, it holds policy elements and, possibly, other policy sets. We denote access requests according to the following notation. Let S, O, A and X denote all subjects, objects, actions and context variables in an access control system respectively.

Definition 4: (Access Request) An access request q is the tuple (s, o, a, x), where $s \in S$ is the subject making the request, $o \in O$ is the requested object, $a \in A$ is the requested action on object o, and $x \in X$ are the context attributes.

```
<PolicySet PolicySetId="PSID"
   PolicyCombiningAlgId="permit-overrides">
   <Target/>
   <Policy PolicyId="PID"
      RuleCombiningAlgId="permit-overrides">
      <Target/>
      <Rule RuleId="RID1" Effect="Deny">
         <Target>
          <Subjects>
            <Subject>Bob</Subject>
            <Subject>John</Subject>
          </Subjects>
          <Resources>
            <Resource>file2</Resource>
          </Resources>
          <Actions>
            <Action>
              <ActionMatch MatchId="string-equal">
                <AttributeValue DataType ="string">
                   read
                </AttributeValue>
                <ActionAttributeDesignator
                 AttributeId ="AID1" DataType ="string"/>
              </ActionMatch>
            </Action>
          </Actions>
         </Target>
      </Rule>
      <Rule RuleId="RID2" Effect="Permit">
         <Target>
          <Subjects>
            <Subject>Bob</Subject>
          </Subjects>
          <Resources>
            <Resource>file1</Resource>
          </Resources>
          <Actions>
            <Action>
              <ActionMatch MatchId="string-equal">
                <AttributeValue DataType ="string">
                   read
                </AttributeValue>
                <ActionAttributeDesignator
                 AttributeId ="AID2" DataType ="string"/>
              </ActionMatch>
            </Action>
          </Actions>
         </Target>
      </Rule>
   </Policy>
</PolicySet>
```

Fig. 1. XACML Policy Set example.

Let us consider the PolicySet listed in Figure 1 which contains one policy with 2 rules. The first rule specifies that "Both Bob and John are denied read access to file2" where each "Bob" and "John" is a *Subject*, "denied" is the rule *Effect*, "read" is the *Action*, and "file2" is the *Object or Resource*, whereas the second rule says "Bob has permission to read file1", "Bob" being the *Subject*, "has permission" the *Effect*, "read" the *Action*, and "file1" the *Object*. Either rule could be accompanied with context parameters (Environment Attributes) as part of a rule's condition such as time, system variables, history, or location. A target is a condition on subject $s \in S$, object $o \in O$ and the action $a \in A$. If the request satisfies the target conditions of a rule (policy) then we say that the rule (policy) is *applicable* to the request, otherwise it is *not applicable*. That is, if Bob makes a request to read file1, his request would be applicable to the second rule which would return a Permit.

III. Policy and Rule Reordering Framework

When a web server needs to enforce an XACML policy with a large number of rules, its XACML policy evaluation engine may easily become the performance bottleneck for the server. To enable an XACML policy evaluation engine to process simultaneous requests of large quantities in real time, especially in face of a burst volume of requests, an efficient XACML policy evaluation engine is necessary. In such environments the requests' distribution is dynamic in terms of volume, types and type of requesters. Motivated by such observation, we develop an adaptive framework that dynamically determines the best ordering according to the incoming requests and the recently received history of requests and executions. In this section we present the basic notions that are relevant for our framework, define statistics extracted from policy execution logs, formulate the rule ordering problem, and finally provide an algorithm to provide the optimal rule ordering.

A. Execution Vector and Policy Permutation

In what follows for the sake of presentation we focus on policy permutation where a similar approach is adopted for PolicySet permutation. We define a policy permutation as follows:

Definition 5: (Policy Permutation) Given a policy P with a rule set $P.R = \{r_1, \ldots, r_n\}$, a policy permutation π is a policy P_π generated by the following procedure:

- (0) $P_\pi.R = \{\}$, $P_\pi.id = P.id$, $P_\pi.t = P.t$, and $P_\pi.RC = P.RC$.
- (1) P' is a copy of P.
- (2) Select a random rule r_i from P' and append r_i to the end of P_π.
- (3) Repeat step 2 until P' is empty.

Policy permutation may alter the correctness of a policy, and result in different evaluations for a same set of requests. We are interested in policy permutations that do not alter the policy evaluation results for any request.

Definition 6: (Safe Policy Permutation) A safe policy permutation π of a policy P is safe iff all requests permitted (denied) by the permuted policy P_π are also permitted (denied) by P.

We assume all requests are well formed such that the policy evaluation returns PERMIT or DENY by the PDP. Using such an assumption, we provide the below theorem:

Theorem 1: Safe Permit (Deny) Overrides Permutation. A policy P having a rule combining algorithm $P.RC$ set to Permit-Overrides or Deny-Overrides is safe with respect to all possible policy permutations.

Proof: The semantics of the permit overrides is that if any rule evaluates to permit then the final authorization decision is permit. Assuming each rule returns either permit or deny then the policy evaluation of a policy P, with a permit overrides rule combing algorithm is the disjunction of all the rule results represented by: $E(P) = E(r_1) \vee \cdots \vee E(r_n)$. The disjunction operator is commutative where $a \vee b = b \vee a$, and associative where $(a \vee b) \vee c = a \vee (b \vee c)$, thus the evaluation of the policy P and any permutation P_π are equal $E(P) = E(P_\pi)$. The deny override follows similar semantics and follows a similar proof. ■

Using Theorem 1 policies with permit override or deny override rule combining algorithms can be permuted without affecting the policy semantics. This does not hold for other rule combining algorithms such as First-Applicable. We focus our discussion on permit and deny override combining algorithms for reordering optimization. As discussed in the following sections, policy based categorization is independent of the rule combining algorithm used.

Given a policy permutation π and a given request q, a subset of rules is of relevance. We represent an ordering of such rules as the *execution vector*.

Definition 7: (Execution vector) $\Gamma = [r_1, \ldots, r_n]$ is the execution vector representing the set of applicable rules, where rule r_i is executed before rule r_{i+1}. $\pi(i)$ refers to the position for rule r_i in execution vector.

According to Theorem 1, any policy execution vector for a policy P having permit overrides rule combining algorithm will evaluate to the same effect as P, the challenge is to evaluate the execution vector that will provide the lowest latency. Hence, we need to define the rule weights in order to present our optimal rule ordering approach.

B. Computation of Rule Weights

Our approach relies on statistics and metrics collected as PDP receives requests. Statistics are collected at two separate levels: *policy* and *rule* level. At the policy level, we are interested in understanding how often a policy applies, and by which class of users. At the rule level, it is important to identify the class of efficient execution vectors. In order to collect meaningful metrics, we assign to each rule (policy) weights that reflect the dominance of this rule in the requests. The weights are based on the PDP returned values, and constructed based on the 1) frequency and the 2) complexity of the rule (policy).

During a given time interval the number of times a policy P_i or a rule r_j gets evaluated is referred to as the hit frequency. We refer to the hit frequency by f and use the dot notation to refer to policy $(P_i.f)$ and rule $(r_j.f)$ hit frequency. Statistics with respect to the hit frequency are accumulated as follows:

- *Policy (Rule) Permit Ratio:* Records the ratio between the number of times a policy (rule) returns a permit with respect to the number of times a policy (rule) gets evaluated, where $P_i.p$ and $r_j.p$ represent the policy and rule permit ratios respectively.
- *Policy (Rule) Deny Ratio:* Records the ratio between the number of times a policy (rule) returns a deny with respect to the number of times a policy (rule) gets evaluated. Where $P_i.d$ and $r_j.d$ represent the policy and rule deny ratios respectively.
- *Policy (Rule) Hit Ratio:* Records the ratio between the number of times a policy (rule) is applicable with respect to the number of times a policy (rule) gets evaluated. Where $P_i.a$ and $r_j.a$ represent the policy and rule hit ratios respectively.

Note that all the above statistics are easily derived from the XACML execution log (see Figure 2). In addition to the rule evaluation statistics we also consider the rule computational complexity. Rules vary from simple conditions to more complicated statements that require the parsing of an XML document or querying a database. The rule complexity metric is related to the number of operations required to execute the rule, we compute

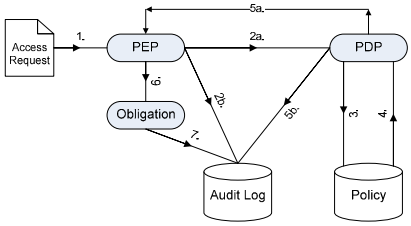


Fig. 2. Log Based XACML Policy Evaluation Framework.

it as the number of boolean atomic conditions appearing in a rule, both at target and at the condition element. Let $n(t)$ denote the number of conditions in the Target element (denoted as t according to Def. 3), and let $n(c)$ be the number of conditions in the Condition element c. XACML supports over 100 standard functions that could be used in the boolean conditions, for example the $Belong\_to$. We assign a cost m_i to each standard function std_i appearing in the rule. m_i is computed by estimating the average execution time of the function. The simple atomic boolean conditions are assigned a constant cost k. For a rule r_j the complexity metric is given by:

$$E_j = k * n(r_j.t) + n(r_j.c) + \sum_{std_i \in r_j} m_i$$

where std_i represents a uniquely identified standard function appearing in r_j. Using both the accumulated rule statistics and the complexity metric for a rule r_j we compute the rule cost as follows:

$$c_j = \beta * E_j + \alpha * F_j$$

Here, β and α are weights that allow system administrators to tune the computation cost, based on the local constraints, such as the available processing power and network bandwidth.

The rule cost is designed to represent the cost of computing a rule, the complexity metric E_j easily represents the rule cost, however the other component is based on the rule's accumulated statistics F_j. The value of F_j is based on the rule combining algorithm, for example if a rule combining algorithm is Permit-Overrides then the metric F_j is based on the decreasing function with respect to the rule permit ratio $(r_j.p)$ or an increasing function with respect to the rule deny ratio $(r_j.d)$. Intuitively, this implies that the rules need to be reordered such that for a policy with the permit overrides rule combining algorithm, the rule r_j with the lowest c_j is to be evaluated first.

C. Optimal Rule Reordering

Using the rule cost metrics we present our optimal rule reordering problem. Given a policy (P_i), the optimal request execution problem (REP) is to find an execution sequence that requires the minimum number of rule evaluations. We assume that rules within policies are evaluated sequentially. The policy P_i, composed of n rules $\{r_1, \ldots, r_n\}$, where $\pi(j)$ refers to the position (depth) for rule r_j in the policy execution vector. The cost associated with rule r_j as computed in Section III-B is referred to as c_j. The expected cost (i.e., average search length) for a given permutation π is given by:

$$\Phi_i = \sum_{j=1}^{n} c_j \pi(j)$$

The main challenge is to compute the optimal policy permutation π that will generate the minimum expected policy execution cost. Additionally, among the possibly optimal π, we need to ensure the policy permutation to be safe, as defined in Definition 6. By computing Φ_i we are able to generate a cost metric for each policy P_i.

A policy set PS is composed of a set of policies $\{P_1, \ldots, P_m\}$. We assume the policies are executed sequentially. Using the minimum policy expected cost Φ_i, and the collected policy evaluation statistics, we compute the policy set execution sequence. The position of policy P_i in the policy set execution sequence is referred to by $\xi(i)$. The expected cost (average search length) for a given policy set (PS_k) permutation ξ is given by:

$$\Psi_k = \sum_{i=1}^{m} \Phi_i \xi(i)$$

The costs Φ_i and Ψ_k are minimized when policies and rules are ordered in ascending order with respect to their costs [17]. Figure 3, shows the algorithm used at both the policyset and policy levels.

```
Algorithm: optimize_policyset
Input: Policy Set PS = {P_1, ..., P_m},
Output: Optimal Policy Set Permutation PS*
 1:   if PS.PC = Permit-Overrides or Deny-Overrides
 2:     PS* ← {}
 3:     for each P_i ∈ PS
 4:       P_i* ← optimize_policy(P_i)
 5:       if PS.alg = Permit-Overrides
 6:         P_i*.c = α * P_i*.Φ + β * P_i.p^-1
 7:       elseif PS.alg = Deny-Overrides
 8:         P_i*.c = α * P_i*.Φ + β * P_i.d^-1
 9:       PS*.insert(P_i*) //Priority Queue on P_i*.c
10:     return PS*
11:   return PS

Algorithm: optimize_policy
Input: Policy P = {r_1, ..., r_n},
Output: Optimal Policy Permutation P*
 1:   if P.RC = Permit-Overrides or Deny-Overrides
 2:     P* ← {}
 3:     for each r_j ∈ P
 4:       E_j = k * (n(r_j.t + r_j.c)) + Σ_{std_i ∈ r_j} m_i
 5:       if P.RC = Permit-Overrides
 6:         F_j = r_j.p^-1
 7:       elseif P.RC = Deny-Overrides
 8:         F_j = r_j.d^-1
 9:       c_j = β * E_j + α * F_j
10:       P*.insert(r_j) //Priority Queue on c_j
11:     P*.Φ = Σ_{j=1}^{n} c_j π(j)
12:     return P*
13:   return P
```

Fig. 3. Optimal PolicySet and Policy Reordering

For example, consider a school database. During certain time periods, the access requests would be more uniform and from the same class of users (e.g. at the beginning of a semester most requests would be from students needing to register for courses, whereas faculty requests will be much less), while during other time periods, more heterogeneous set of requests may be submitted. In section V of this paper, we show how our framework adapts to the different types of requests received and how we can benefit from policy/rule reordering.

Weights can be updated according to two different strategies: 1) periodically, 2) based on the last ρ received requests. In the first case, we update the weight values using the latest statistics. New execution vectors are constructed using fresh rule weights in order to boost up the hit performance close to its optimum level. The update period should be based on the predictable incoming request (e.g., certain months of the year) flow changes. In the latter case, the optimal execution vectors are constructed based on the computed rule weights. The incoming access requests are then processed according to the ordering determined. Intuitively, the maximal reduction is obtained when the incoming requests perfectly match the requests' distribution. Notice that more than one execution vector could be optimal and safe. However, since not all rules have the same complexity, different execution vectors may sensibly influence the overall evaluation time, even if a safe and efficient policy permutation is found.

IV. Categorization Based Optimization

The optimization problem minimizes the average request evaluation time. This approach is ideal if the policy requests follow a uniform statistic. However, this approach is unlikely to be satisfactory in scenarios where the requests' distribution is dynamic in terms of volume and type of requesters. If we solely rely on reordering, assuming a role based access control (RBAC) system of two roles, say *student* and *faculty*, where there are on average 100 student requests for every faculty request, the computed statistics will be guided by the student requests. As such, the optimization problem presented above will favor the student role. Reordering rules and policies in these circumstances is not sufficient, as the computational cost will not be given by the evaluation of the rules themselves, rather it will heavily depend on the time spent on finding the applicable policies to the given request.

Hence, in order to further improve the efficiency of the rule reordering, we resort to clustering the policies. Building on execution vectors, an intuitive mechanism is to categorize the policies based on the subjects. Starting from a set of $L[S]$ clusters, where $L[S]$ is the number of subjects in S, the goal is first to reduce the number of categories in order to allow the reordering to have a considerable effect on the execution time. Second, to reduce the memory footprint needed for caching the categories. When the categorization is done on a per-subject basis, to record an improvement in the execution time the policies must be adequately large. This happens because, when there is a category for each subject, there is essentially a unique execution vector for that subject. When large policies are evaluated, the categorization helps provide a good match for the execution vector and hence fewer rules are evaluated, thereby improving the evaluation time. In case of small policies, to make categorization effective, we need to decrease the number of categories to be searched in order to find the execution vector. In order to resolve this issue, we resort to further clustering the requests. Figure 4, shows a PolicySet and the different applicable views based on the involved subject, where each view could serve as a subject based category.

To achieve these results, we propose adopting an algorithm based on the K-Means clustering method [20]. Generally speaking, the K-Means algorithm is used to cluster m objects based on attributes into k partitions, $k < m$. Each cluster consists of a "center" around which individual elements of the data set being clustered are grouped together. This grouping is done based on some measure of similarity to the other elements in that cluster. In our domain, the number of clusters N_c and the centers of these clusters, i.e. N_c subjects are chosen at random from the set of subjects S. The set of centers (or clusters) is referred to as C_s. Each subject $S_i \in S$ is considered, and its similarity $D_{i,k}$ is calculated with respect to each subject $S_k \in C_s$ in the different

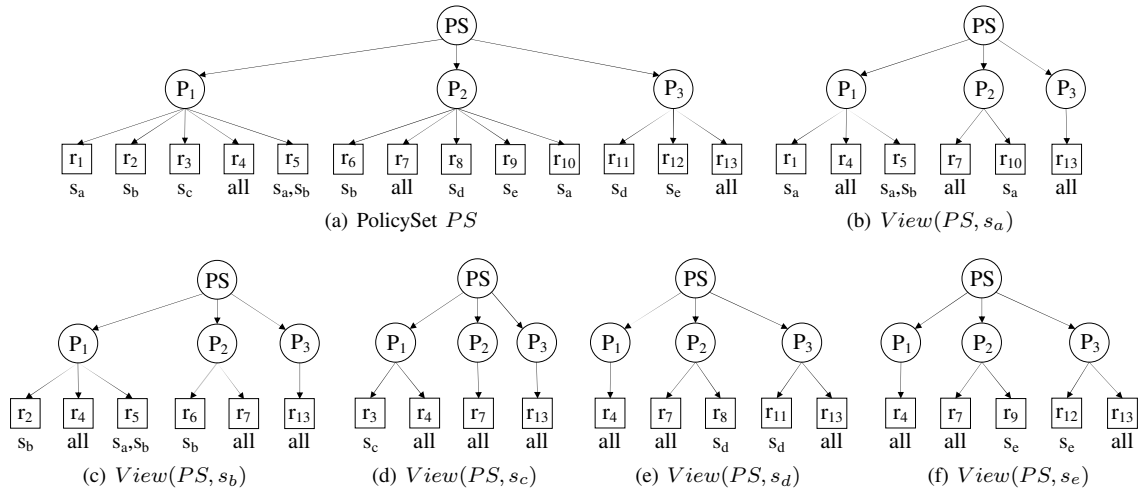


Fig. 4. Policy Set and Views

clusters. S_i will be added to that cluster where the similarity $D_{i,k}$ is maximum. The strength of this simple algorithm lies in the way the similarity metric $D_{i,k}$ is calculated. The similarity metric aims to cluster together the subjects that share a large number of policies which are applicable to all of them. Let $\mathbb{P}_i$ represents the set of policies applicable to a given subject S_i and let $L[\mathbb{P}_i]$ be the number of policies applicable to that subject. The number of policies shared between two subjects, S_i and S_k is given by $L[\mathbb{P}_i \cap \mathbb{P}_k]$. The fraction of the number of policies shared between the two subjects that are a part of $L[\mathbb{P}_i]$ is given by $\Theta_{i,k}$, where:

$$\Theta_{i,k} = \frac{L[\mathbb{P}_i \cap \mathbb{P}_k]}{L[\mathbb{P}_i]}$$

The similarity metric $D_{i,k}$ between subject S_i and S_k is calculated as follows $D_{i,k} = \Theta_{i,k} + \Theta_{k,i}$. The subject S_i is grouped with the cluster centering on S_k where $D_{i,k}$ is maximum. This ensures that only those subjects which have a large number of policies in common are grouped together. In general, the clustering is more effective when the number of shared policies is large, i.e. when $L[\mathbb{P}_i \cap \mathbb{P}_k]$ is large. The number of clusters N_c should be chosen carefully. The larger the value N_c, the lesser visible will the effect of reordering be. This is more evident when we consider the fact that as N_c approaches $L[S]$, we essentially experience the initial effect of having $L[S]$ unique categories for each of the subjects. On the other hand, should N_c be too small, the improvement obtained by categorization is completely lost, because as N_c approaches '1', all the subjects belong to the same cluster. In other words, there are no clusters at all.

This algorithm allows us to tune our optimization approach such that we can either maximize the improvement due to clustering or due to reordering, or both, based on the specific context. In general, the improvement due to clustering and categorization is most apparent when there are very large policies to process. On the other hand, for extremely simple policies with only one or two subjects, reordering is more helpful. In this scenario, reordering saves valuable execution time because by reordering, we can ensure that the policy does not do a brute force search to evaluate all the rules.

V. Experimental Results

Our experiments were performed on a MacBook Pro running Mac OS X 10.5.5 with 4GB of RAM and a 2.4GHz Dual Core Intel processor. Experiments were conducted on both synthetic policies and real world-based policies. The synthetic polices were divided into two sets of test suites. The first test suite deals with XACML policy sets where subjects have a small number of applicable rules. The second suite investigates policy sets where subjects have a large number of applicable rules, and will show the significant effect of applying our reordering technique to large policy sets. In other words reordering will not have a big effect if on a policy with only 2 rules, whereas it will make difference with a policy that has say 100 rules or more.

The real world-based policy sets are policies built using existing data sets, and properly modified to fit our framework without changing the semantics -or the structure- of the policies. Precisely, we tested the policies by Fisler et al. [4], which they used for their Margrave tool. Our experiments ensure that all policies are loaded into memory before executing any request evaluations. This ensures that evaluation times are not skewed by any policy loading time. All tests were conducted using 100,000 randomly generated XACML requests. All requests have a single value for the subject, resource, and action. Using the real world-based policy sets and the synthetic test suites, we performed extensive experiments to investigate the performance enhancements yielded by our proposed categorization and reordering techniques. We also compared our results with Sun's PDP engine results.

Our experimental process includes two main stages; first, the setup stage and second, the actual request evaluations. The setup stage includes three sub-stages:

S1. Categorization of the experimental policy sets. Categorization is performed as explained in Section IV. The number of categories used for each policy set ranges from N to $N/10$, where N is the number of unique subjects within a policy set,

S2. Training stage that collects the results of request evaluations (permit, deny, not-applicable, indeterminate) subsequently used for the reordering stage,

S3. Reordering policies within the policy set and all rules within each policy according to the statistics we gathered during the

training stage.

The setup stage needs to be executed only once, however the sub-stages (S2) and (S3) could be executed repeatedly to retrain and reorder the policies and rules to achieve better performance. For our tests, we chose not to repeat the sub-stages, and thus measure the performance in the worst case scenario. The results of categorization and reordering are cached in memory. During the second stage the access requests are actually evaluated, using the ordering and categories set up in the previous stage. The processing time is the time needed to evaluate a request against a policy subjected to our setup stage plus the time to make a decision on that request. The preprocessing time is the time needed to complete the setup stage.

When designing our framework we tried to introduce as few changes as possible to Sun's PDP for easy integration with existing Sun PDPs. Our framework does not introduce a new method for evaluating a request, the matter of fact is we use the core of Suns PDP to do request evaluations. Hence, attribute retrieval times apply to both our framework and Sun's PDP equally, and are considered as part of the total evaluation time of a request. Our framework benefits from categorizing and reordering policies/rules, which results in avoiding any unnecessary evaluations and evaluating policies/rules that result in faster decisions, rather than introducing a new method for evaluating requests.

The experimental results show that our framework is orders of magnitude more efficient than Sun's PDP, and the performance difference between our framework and Sun's PDP grows linearly with the number of requests and number of rules within a policy set. We discuss the test results in detail in the following subsections.

A. Real World-Based policies

The experiments on real world-based policies used the policy sets by Fisler et al. [4], specifically CodeA, CodeB, CodeC, & CodeD. We also added another policy that we call CodeDMod, which is an enlarged version of the policy CodeD. This policy set contains 11 policies and 75 rules in total. We include this policy in order to evaluate the performance of our framework with larger real world policies. As highlighted by [8], it is difficult to access large real world policies that are publicly available, due to the confidential information these policies typically carry. Another issue highlighted by other authors [7] is the fact that XACML policies tend to get larger and more complicated with time, hence we introduced CodeDMod to represent such a large policy.

Table I summarizes the results of the experiments done on the real world-based policies. In all cases we obtain at least a 78% performance improvement over Sun's PDP. Despite the nature of our framework which best suits large policies, our optimization engine still provides a significant performance boost in the case of smaller policies, e.g. CodeA is a policy set with only 2 rules. The policy CodeDMod which is a much larger policy, shows a performance boost of over 91% over Sun's PDP. We also notice the difference between using categorization only and the effect of adding reordering to the framework. Reordering boosts the evaluation performance up to 22% over using categorization only. This is noticeable in the case of CodeDMod where reordering has an effect on its 11 policies' and 75 rules' order. In the smaller policy sets CodeA, CodeB, CodeC, & CodeD, reordering does not provide a big performance boost over categorization only, but still gives up to 8.5% better performance in the case of CodeA.

B. Synthetic Policies

We test our framework against large synthetic policies to show the scalability of the framework and the high performance that it provides in the case of very large policies. We divide the synthetic policies into two test suites, each of which has policy sets of sizes ranging from 400 to 4000 rules. The following sections explain the test suites results in detail.

1) Test Suite I Results: This test suite deals with policy sets where each subject has a few number of applicable rules. This test case is used to emphasize the effect of our categorization technique, whereas our reordering technique may have a minor effect. This test suite uses policy sets of 4000, 2000, 1000, and 400 rules. For each policy set, rules are divided evenly among 100 policies. For the sake of testing the *Permit Overrides* combining algorithm is used for all the test policy sets and policies. Using this test suite our approach is 1638 times faster than the Sun PDP.

a) Results with Categorization Only: We carried out a first set of tests only applying the categorization technique with no reordering. The number of categories used for each policy set was varied from N to $N/10$, where N is the number of unique subjects within a policy set. The preprocessing time for this approach is the time needed for categorizing a policy set (sub-stage S1.). When using N categories, results show that preprocessing a policy set of 100 policies and 4000 rules takes about 25138 ms and a policy set of 100 policies and 400 rules takes about 913 ms. When $N/10$ categories are used, preprocessing times are 23464 ms and 487 ms for the 4000-rule and 400-rule policy sets respectively.

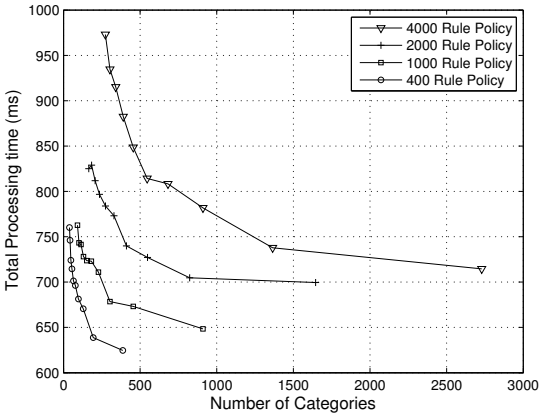
The experimental results demonstrate that the total processing times for our approach is at least 172 times faster than Sun's PDP. For a policy set of 100 policies and 4000 rules while using $N/10$ categories, it takes 973.1 ms to evaluate 100,000 random requests, whereas Sun's PDP takes about 1152460 ms. A policy set with 400 rules takes 760.2 ms and Sun's PDP takes about 130421.3 ms. When N categories are used, total processing times are 714.6 ms and 624.6 ms for the 4000-rule and 400-rule policy sets respectively. Figure 5(a) shows the complete results when using categorization alone with respect to the number of categories used, which range from 0 to 3000.

b) Results with Categorization plus Reordering: For this set of tests, we applied the categorization technique, followed by our reordering technique. The number of categories used also range from N to $N/10$. We make use of all sub-stages within the setup stage. Preprocessing time in this case is the time for both categorization and reordering of rules. The results for this set of tests are reported in Figure 5(b). The experimental results shows that the total processing times for our approach is at least 171 times faster than Sun's PDP. For a policy set of 100 policies and 4000 rules while using $N/10$ categories, it takes 967.5 ms to evaluate 100,000 random requests, whereas Sun's PDP takes about 1152460 ms. A policy set with 400 rules takes 763 ms and Sun's PDP takes about 130421.3 ms. When N categories are used, total processing times are 703.7 ms and 616.2 ms for the 4000-rule and 400-rule policy sets respectively. Figure 5(b) shows our complete results when using categorization plus reordering with respect to the number of categories used. Figure 5(c) is a comparison between our approach with categorization plus reordering and Sun's PDP. The plot representing our approach is an average of the best and worst case we obtained from using

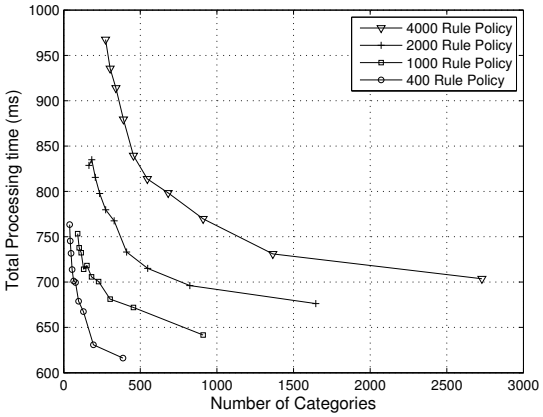
Policy	#Rules	Sun	Categorization	Categorization & Reordering	Cat.-Only Boost	Cat.+Reordering Boost
CodeA	2	867	152	139	82.47%	83.97%
CodeB	3	1007	191.3	191	81%	81.3%
CodeC	4	1007.5	200.1	200	80.14%	80.15%
CodeD	5	1150	249	242.8	78.35%	78.89%
CodeDMod	75	32223	3474	2709	89.22%	91.59%

TABLE I

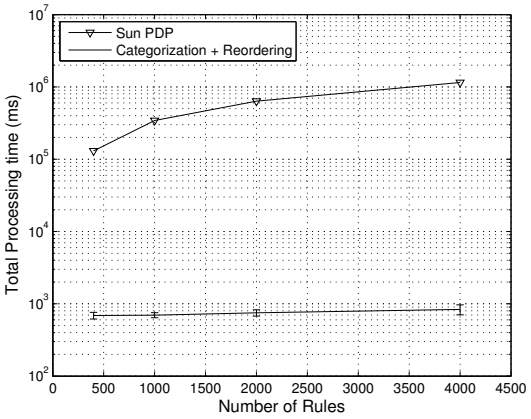
REAL WORLD-BASED EVALUATION RESULTS IN MILLISECONDS



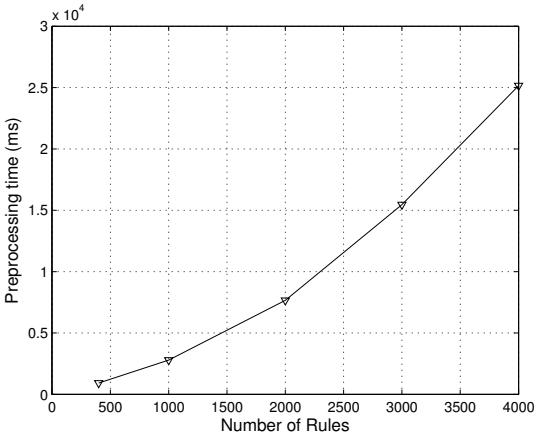
(a) Effect of categorization on evaluation time w.r.t # of categories used with no reordering.



(b) Effect of categorization and reordering on evaluation time w.r.t # categories used.



(c) Evaluation times comparison between our approach and Sun PDP.



(d) Preprocessing times including categorization and reordering.

Fig. 5. Experimental Results for Test Suite 1.

different numbers of categories. The results obtained by this set of tests report a very slight performance improvement due to the reordering.

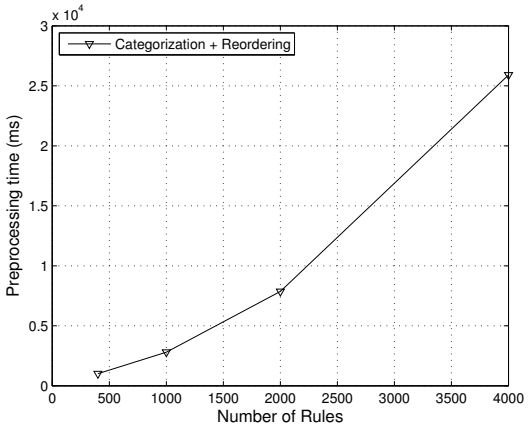
Reordering rules is not a significant factor to performance because of the low number of rules applicable to each subject. Reordering's effect can be better appreciated for policy sets with many rules applicable to each subject.

With regards to preprocessing, our results show that preprocessing time is proportional to the number of rules, as reported in Figure 5(d). Preprocessing a policy set of 100 policies and 4000 rules while using N categories takes about 25158 ms, and a policy set with 100 policies and 400 rules takes about 925 ms. When $N/10$ categories are used, preprocessing times are lower, 23472 ms and 491 ms for the 4000-rule and 400-rule policy sets respectively. Our tests also show that the preprocessing times are proportional to the number of categories used. More categories lead to higher preprocessing times due to the extra processing needed to match similar subjects to a common category. Next, we present a second test suite highlighting the advantages of the reordering effect.

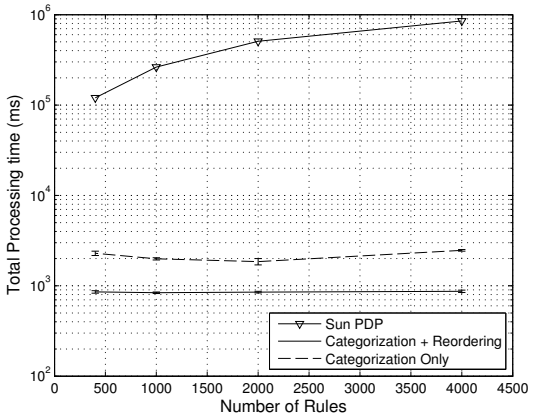
2) Test Suite II Results: Our first test suite did not give us any indications about the effect of reordering on the policy evaluation performance.

This is due to the fact that most subjects had a small number of applicable rules, which led to subjects having very small policy/rule execution vectors. As a result, reordering only happened on small execution vectors, and therefore did not make a significant difference.

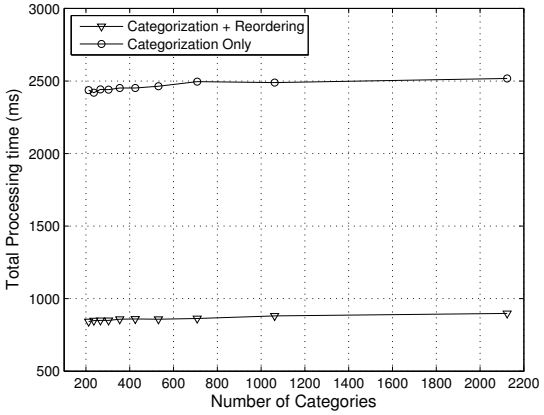
Hence, we decided to generate a second test suite that could



(a) Preprocessing times (categorization + reordering).



(b) Sun PDP evaluation times compared to categorization only and categorization + reordering.



(c) Performance boost from reordering w.r.t the # of categories using a 4000-rule policy.

Fig. 6. Experimental Results for Test Suite 2.

allow us to observe the impact of reordering on performance. This suite simulates a scenario where each subject within a policy set is guaranteed to have a significant number of applicable rules. This case might occur when a specific subject has high privileges and has access to a high number of resources. In this case the subject will have a high number of rules permitting him access to these resources.

When reordering happens in such a scenario, there will be no need to go over all rules within a subject's category. As expected, this test suite showed a significant performance advantage for the categorization plus reordering approach over the categorization only approach. We used policy sets of 4000, 2000, 1000, and 400 rules (different from the ones used in first test suite). For each policy set, rules are divided evenly among 100 policies. Overall, our results for this test suite show that our approach is 949 times faster than Sun's PDP engine. Similar to the first test suite, we conducted experiments using categorization only and categorization with reordering.

a) Results with Categorization Only: The preprocessing times for this case are inline with the times for the analogous set of tests (Section V-B.1) of the first test suite. Precisely, when using N categories, preprocessing a policy set of 100 policies and 4000 rules takes about 25397 ms and a policy set of 100 policies and 400 rules takes about 978 ms. When $N/10$ categories are used, preprocessing times are 28633 ms and 1075 ms for the 4000-rule and 400-rule policy sets respectively.

As in the previous test case, the results for total processing times show a very significant improvement in performance over Sun's PDP. Our results indicate that our mechanism provides at least 48 times faster evaluation. For a policy set of 100 policies and 4000 rules while using $N/10$ categories, it takes 2437.2 ms to evaluate 100,000 random requests, whereas Sun's PDP takes about 851477 ms. A policy set with 400 rules takes 2272.2 ms and Sun's PDP takes about 120230.3 ms. For N categories, total processing times are 2517.6 ms and 2242.5 ms for the 4000-rule and 400-rule policy sets respectively.

b) Results with Categorization plus Reordering: Figure 6(a) reports the preprocessing times for this approach. Our results show that preprocessing a policy set of 100 policies and 4000 rules while using N categories takes about 25902 ms and a policy set with 100 policies and 400 rules takes about 1007 ms. When $N/10$ categories are used, preprocessing times are 31052 ms and 1061 ms for the 4000-rule and 400-rule policy sets respectively. Although the policies are different, we notice that the gathered

times are very similar to the times recorded for preprocessing the set of policies used for the first test suite (reported in Figure 5(d)). This observation leads to the conclusion that the preprocessing time is not influenced by the type of policies used. The preprocessing times are almost negligible when compared to the highly significant performance improvement in total processing times over Sun's PDP, not to mention that preprocessing times correspond to the setup stage of our framework which only occurs once within a policy set's lifetime or upon a client's request.

Figure 6(b) compares Sun's PDP total evaluation times with our results from the second test suite. The total processing time of our approach is at least 139 times faster than Sun's PDP. As shown, for a policy set of 100 policies and 4000 rules while using $N/10$ categories, it takes 842.3 ms to evaluate 100,000 random requests, whereas Sun's PDP takes about 851477 ms. A policy set with 400 rules takes 867.5 ms and Sun's PDP takes about 120230.3 ms. When N categories are used, total processing times are 897.6 ms and 830 ms for the 4000-rule and 400-rule policy sets respectively.

For the 4000-rule policy set used in this test suite, results indicate that categorization plus reordering has a 65.4% performance improvement over using categorization alone. Figure 6(c) shows the performance boost reordering provides with respect to the number of categorizations used. The figure shows that adding reordering to categorization provides over 1.6 seconds of an advantage over the use of categorization only.

We notice a slight improvement in performance when the number of categories is reduced. This result is explained by the fact that the policy set we used has many rules that are applicable to all subjects, which means the resulting categories are not much different from the original categories.

C. Adaptability of Reordering Approach

Figure 7, demonstrates how our reordering approach adapts to the incoming requests received by the PDP. As mentioned earlier in the reordering approach, we have a reordering process that reorders both policies within a PolicySet and rules within all policies. The reordering happens according to the number of Permits/Denies a policy or rule triggers. Figure 7 shows how the order of 10 policies within a PolicySet changes with respect to time. The orders of policies ranges from 0 to $L[P_i]$, where $L[P_i]$ is the number of applicable policies for subject S_i (The size of a subject's policy execution vector). Order 0 reflects the highest ranked policy (the policy most requested). Figure 7 shows the policies within a policy execution vector for a particular subject, in this case subject S_1. It is important to notice that each reordering cycle (a single reordering process) is dependent on all previous cycles. In Figure 7, t_0 represents the initial time before reordering, and t_n represents the time at which n reordering cycles have been executed (reordering of policies/rules based on the evaluation results at $t_{n-1}, t_{n-2}, \ldots, t_0$). As time passes and more reordering cycles occur, one can notice how the order of some policies starts to settle at a certain position. For instance, if one looks at policy P_7, it gets pushed to order 9 at t_1, this is due to the low number of Permits/Denies returned by this policy. Whereas if one looks at policy P_0, it gets to order 1 and stays there as it is requested very frequently. Policy P_4 settles after t_7. Other polices settle for a while and then get reordered as the incoming requests might influence their order positions. The ordering of these policies depends on the incoming requests and how they trigger the accumulated number of Permits/Denies a policy evaluates to. Each subject within a policy set will reflect a similar adaptation process to the one in Figure 7, each of which prioritizes their applicable policies and rules according to the statistics from previous reordering cycles.

To clarify how the adaptation process would actually occur, let us look into a case scenario e.g. a school. At the beginning of a semester, most access requests would be driven by students wanting to register for their courses. The adaptation process would move policies/rules that are applicable to students and favor their incoming requests to the top of a policy set, which will result in faster evaluation times for such similar future requests. Within a semester, where most midterms are given, many faculty requests for inserting or updating student grades will be recorded. In this case, the adaptation process will favor faculty requests by moving policies/rules within a policy set to the top, and hence favoring these requests. Whenever there is a flow of similar requests from different subjects within the school, the policy set will adapt to the best configuration that will result in the best evaluation results.

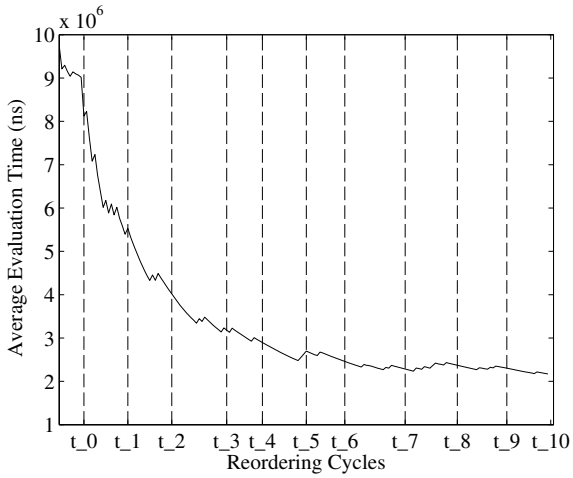


Fig. 8. Average Request Evaluation Time and Reordering Cycles.

Figure 8, demonstrates the average request evaluation times for subject S_1 with respect to time. As the time proceeds, a number of reordering cycles occur, hence influencing the order of subject S_1's policies within its policy execution vector and rules within its rule execution vector. The reordering process will push the most requested policies and rules that evaluate to Permit/Deny up to the front of the corresponding execution vectors. This will result in faster evaluation times as depicted by our test results in Figure 8. Note that the average request evaluation time gradually decreases as more reordering cycles are executed and thus adapt to the incoming different request trends.

Space Complexity. Concerning space complexity, our framework is relatively efficient. After sub-stage *a*, the categorized policy set will be cached in memory using a Hashtable (H_1). H_1 will be of size $N_c * L[\mathbb{P}_c] * L[\mathbb{R}_c]$, where N_c is the number of categories used, $L[\mathbb{P}_c]$ the number of policies within a category, and $L[\mathbb{R}_c]$ the number of rules within $\mathbb{P}_c$.

In section VI, we highlight some related work, and compare their results with ours where applicable.

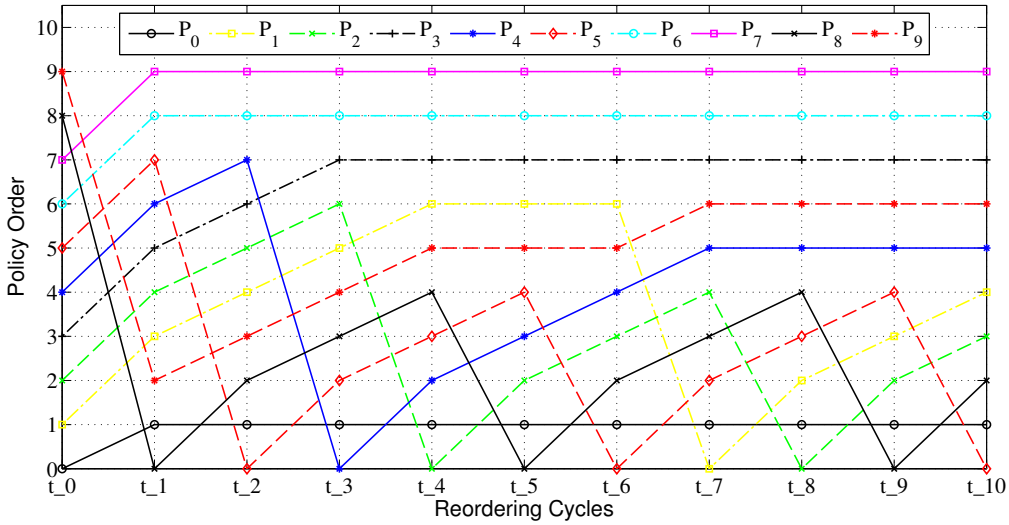


Fig. 7. Policy Order and Reordering Cycles

VI. Related work

XACML 1.0 had been standardized by the OASIS Committee in 2003 and ever since a lot work has been done on the analysis, testing, verification and optimization of XACML Policies [13], [10], [11], [2], [18]. Many have recently focused on issues related to XACML optimization and analysis [11], [14], [2], [18]. In [11], Liu et al. present one of the most interesting proposals on optimization of XACML policies so far. Liu et al, focus on improving the performance of the PDP by numericalization and normalization of the XACML Policies. The numericalization is used to convert the string policies into numbers. The authors posit that since numerical comparison is more efficient, an improvement in performance is achieved by numericalization itself. Further, normalized policies are converted into a flat policy structure. In doing this, the authors replace the different rule-combining algorithms with only on, viz. First-Applicable. They then proceed to convert the numericalized, normalized policies into tree data structures for efficient processing. While this paper differs from our work in the approach used to process the policies, we also focus on reordering of the rules to achieve an improvement in the performance of the PDP. In other words, the underlying concept is the same: the brute force evaluation of a policy by the PDP can be avoided, if the rules are reordered such that the rule which is applicable to the rule combining algorithm is evaluated first and the time for processing or evaluating not applicable rules is saved. We identify two main differences between our work and [11]. First, we rely on statistics, which help us defining the best ordering process based on *actual* user requests. Secondly, the authors unify all the rule combining algorithms into only the First-Applicable, while we do not require such cumbersome preprocessing stage.

Finally, although no complexity evaluation is given by Liu et al., we believe our approach is more efficient in terms of space complexity, since it does not rely on storing complex data structures such as tables and trees.

Another work related to ours is [14], by Miseldine. Miseldine proposes to achieve policy optimization by minimizing the average cost of finding a match at the rule level the target level and the policy level. The work assumes no changes to the XACML specification, in that the Sun's XACML implementation is not altered. Its PDP and PEP engines remain the same. The improvement is achieved by applying optimization techniques to the policies themselves. Therefore, anyone who consumes XACML remains structurally unaffected but anyone who generates XACML policies can generate an optimized output by applying the optimization techniques outlined. The author also builds on the same premise as us that, if a policy or rule is not applicable to a particular target, it needs not be evaluated. The main differences between [14] and our work arise in the way we try to meet this premise. While we focus on the reordering of rules and further on the categorization of the policies based on both the policies and the rules, Miseldine approaches this problem considering *policy configurations*. A policy configuration is the relationship of policy and rule targets to members of the set of rules R, the set of subject S and the set of actions A. Combinations of sets are sought such that policy targets are formed from S.R, R.A or S.A . The match at the rule level must then reference the remaining unused set. Although interesting, the improvements are drastically worse than ours. For example, their optimized method takes around 200ms for evaluating 4000 rules, where with our techniques, it only takes 1 ms. A related approach that has been proposed to improve the performance of XACML evaluation is the recent work on detecting and removing redundant XACML rules [9]. In [9] Kolovski formalizes XACML policies using description logics (DL), and exploit existing DL verifiers to conduct policy verification. Their policy verification framework can detect redundant XACML rules. The idea of removing redundant policies is interesting and may be useful to boost the evaluation time. However, it is yet to be validated whether the improvement will be worth the time needed to remove redundant policies, and how significant the overall improvement would be. If [9] were to be used as a means of enhancing policy evaluations, the preprocessing time (Initial processing + verifying a policy) for the Continue policy (306 rules) used by [4] takes approximately 11 seconds, compared to our framework that takes less than 2 seconds to preprocess a policy set of 400 rules, and

less than 10 seconds for a policy set of 2000 rules.

On the front of testing policies, the work by Martin et al. [12] delves into the problems of defining and measuring policy coverage when testing policies. The techniques presented in [12] are very different from ours. However, in light of the extensive testing needed to ensure that an improvement in performance has indeed been achieved, we drew on the concepts in [12] to ensure that policy coverage has indeed been achieved.

One related area where similar optimization techniques are often explored is Firewall Filtering [5], [6]. In this respect, our work on optimization of XACML policies shares some similarities to the optimization of firewall filtering approaches. The major differences between firewall optimization and XACML policy optimization arise because in the case of firewalls, a major portion of the traffic packets match a small subset of the firewall rules, and the same distribution of traffic is maintained over a significant period of time. This skewness is not experienced in the incoming requests for an XACML policy. Besides, firewall rules, which have dependencies on each other, have an order of precedence defined, while rules in an XACML policies are not related. These two properties of firewall rules allow the authors to prove in [5] and [6], that the optimal firewall rule ordering problem is NP-Complete. Despite these differences between firewall filtering optimization and optimization of XACML policies, we can still draw from the body of work on firewalls, specifically from [5]. We employ metrics similar to the ones used by the authors for evaluating which rules would be most applicable to our policies. Of the different metrics presented, we rely on the frequency of the rules, as in [5]. Frequency is useful in predicting the best match for a new incoming request which does not match any existing categories. The metrics cannot be applied directly in our context, as there are substantial differences between the packet matching algorithm used for the firewalls and the categorization and matching of requests required by our approach. The packet matching is a simple, single level problem as the only requirement is to match the packet's header against the rule list and performing the corresponding filtering. The rule frequency and recency are then updated for the applicable rule. Our goal is more ambitious, since not only we try to find which policy is applicable to an incoming request but also we optimize the ordering of the rules within the policy to match the request. In order to allow for this type of matching, we categorize the requests at two levels. We first categorize the request by subjects to see which policy would be applicable to it and then we further find a match between the request and the execution vectors for that policy. Execution vectors are simply the order in which the rules in a policy are applicable to a request.

Another area related to our approach for policy categorization is packet classification [1]. While similar in spirit to the categorization used by us, there are many differences between the two approaches. In [1] structure of the packet classifiers is flat whereas we need to categorize at the multiple levels of the policies and the rules. Secondly, the rule combining algorithms in firewalls are only first-match or multi-match as opposed to the many different rule combining algorithms for XACML Policies. The approach used for packet classifiers is therefore totally different from the one employed by us for categorization. The approach used in [1] focuses on modifying the rules and reducing the number of entries needed to modify the rules in order to improve packet classification efficiency. We do not need to focus on modifying the rules themselves in XACML policies and our categorization focuses on allowing us to reorder the rules for improved evaluation efficiency by the PDP.

VII. Conclusions and Future Work

XACML policies and their evaluation play a critical role in many access control systems, where numerous requests are received by large set of subjects. This calls for high performance XACML policy evaluation engines. In this paper, we introduced a novel optimization framework based on statistics and policy set categorization. Our categorization technique, which is based on the K-means algorithm, provides fast access to applicable policies and rules for a certain subject. Reordering policies and rules within a policy set ensures that request evaluations are done on policies and rules that are most likely to return a positive effect; hence, avoid examining all policies and rules which are not likely to be significant for the access request being evaluated. We showed through experimental analysis the enhancement obtained for different set of policies of varying size and structures. Our results show that our techniques outperform the policy evaluation of the Sun PDP engine by orders of magnitude. Our framework can be utilized for multiple purposes, besides optimization. For example, our framework could successfully be employed as a debugging or profiling tool for XACML policies. Our current system could be extended so as to provide feedback to policy writers and administrators about the behavior of the system they're authoring, possibly allowing the policy author to determine ordering, rather than automatically set it. Further, as part of our future work, in order to fully assess the potential of our work, we plan on deploying our framework into a real world environment and observe how it affects performance. Since our framework targets very large policy sets, we would like to further investigate the effect of evaluating very small policy sets, e.g., a policy set with a single policy and a single rule. The initial results provided in this direction in Section V-A are encouraging, although additional extensive tests are required, to possibly identify aspects of the system that can be further tuned to improve even in such cases.

Acknowledgment

This work was partially funded by the National Science Foundation (NSF-CNS-0831360) and National Security Agency (NSA H98230-07-1-0231).

References

[1] Q. Dong, S. Banerjee, J. Wang, D. Agrawal, and A. Shukla. Packet classifiers in ternary cams can be smaller. *SIGMETRICS Perform. Eval. Rev.*, 34(1):311–322, 2006.

[2] D. el Diehn I. Abou-Tair, S. Berlik, and U. Kelter. Enforcing privacy by means of an ontology driven xacml framework. In *IAS '07: Proceedings of the Third International Symposium on Information Assurance and Security*, pages 279–284, Washington, DC, USA, 2007. IEEE Computer Society.

[3] Facebook. *http://www.facebook.com*, 2007.

[4] K. Fisler, S. Krishnamurthi, L. A. Meyerovich, and M. C. Tschantz. Verification and change-impact analysis of access-control policies. In *ICSE '05: Proceedings of the 27th international conference on Software engineering*, pages 196–205, New York, NY, USA, 2005. ACM.

[5] H. Hamed and E. Al-Shaer. Dynamic rule-ordering optimization for high-speed firewall filtering. In *Proceedings of the 2006 ACM Symposium on Information, computer and communications security*, pages 332–342, New York, NY, USA, 2006. ACM.

[6] H. Hamed, A. El-Atawy, and E. Al-Shaer. Adaptive statistical optimization techniques for firewall packet filtering. In *INFOCOM 2006: Proceedings of the 25th IEEE International Conference on Computer Communications*, pages 1–12, April 2006.

[7] G. Hughes and T. Bultan. Automated verification of xacml policies using a sat solver. In *Proceedings of the Workshop on Web Quality, Verification and Validation (WQVV 07)*, pages 378–392, 2007.

[8] V. Kolovski and J. Hendler. XACML policy analysis using description logics. *Submitted to Journal of Computer Security available at http://www.mindswap.org/∼kolovski/KolovskiXACMLAnalysis-JCSSubmission.pdf*, 2008.

[9] V. Kolovski, J. Hendler, and B. Parsia. Analyzing web access control policies. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 677–686, New York, NY, USA, 2007. ACM.

[10] D. Lin, P. Rao, E. Bertino, and J. Lobo. An approach to evaluate policy similarity. In *SACMAT '07: Proceedings of the 12th ACM symposium on Access control models and technologies*, pages 1–10, New York, NY, USA, 2007. ACM.

[11] A. X. Liu, F. Chen, J. Hwang, and T. Xie. Xengine: a fast and scalable xacml policy evaluation engine. In *Proceedings of the ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 265–276, New York, NY, USA, 2008. ACM.

[12] E. Martin, T. Xie, and T. Yu. Defining and measuring policy coverage in testing access control policies. In *In Proc. 8th International Conference on Information and Communications Security*, pages 139–158, 2006.

[13] P. Mazzoleni, B. Crispo, S. Sivasubramanian, and E. Bertino. Xacml policy integration algorithms. *ACM Trans. Inf. Syst. Secur.*, 11(1), 2008.

[14] P. L. Miseldine. Automated xacml policy reconfiguration for evaluation optimisation. In *Proceedings of the 4th International Workshop on Software Engineering for Secure Systems*, pages 1–8, New York, NY, USA, 2008. ACM.

[15] T. Moses. Extensible access control markup language (XACML). *Technical Report, OASIS*, 2003.

[16] MySpace. *http://www.myspace.com*, 2007.

[17] R. Rivest. On self-organizing sequential search heuristics. *Commun. ACM*, 19(2):63–67, 1976.

[18] S. Rizvi, A. Mendelzon, S. Sudarshan, and P. Roy. Extending query rewriting techniques for fine-grained access control. In *Proceedings of the International Conference on Management of Data*, pages 551–562, New York, NY, USA, 2004. ACM.

[19] Sun XACML Policy Engine. http://sunxacml.sourceforge.net/guide.html.

[20] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. 2 edition, 2005.

Said Marouf Received a Masters degree in Software Engineering from the University of Wisonsin - La Crosse, in 2008. Currently working towards a PhD degree in Information Technology at the University of North Carolina at Charlotte. Research interests include access control policy optimization & management within social networks and SELinux, secure software development, and vulnerability analysis within the Java programming language execution environment.

Mohamed Shehab is an Assistant Professor of Software and Information Systems Department at the University of North Carolina at Charlotte. He received his PhD degree in Electrical and Computer Engineering from Purdue University in August 2007. His research and teaching interests are in the broad areas of network and information security. In particular, his research focuses on advancing the state of the art in the design and implementation of distributed access control protocols to cope with the requirements of emerging distributed, Web Services, Social Networks, and Peer-to-Peer Environments.

Anna Squicciarini is an assistant professor at the college of Information Science and Technology, at the Pennsylvania State University. During the years of 2006-2007 she was a post doctoral research associate at Purdue University. Squicciarini's main interests include access control for distributed systems, privacy, security for Web 2.0 technologies and grid computing. Squicciarini earned her Ph.D. in Computer Science from the University of Milan, Italy, in February 2006. During her PhD she has been a visiting scholar at the Computer Science Department of Purdue University, at the Colorado State University, and at the Swedish Institute of Computer Science. Squicciarini is the author or co-author of more than 40 refereed journals, and in proceedings of international conferences and symposia. She is an IEEE member.

Smitha Sundareswaran a first year Ph.D student in Penn State's IST Department. She works with Dr. Anna Squicciarini. Her research interests include Web Security and Privacy, Web 2.0 privacy and Digital Identity Management.