Access Control Policy Verification and Repairing in Alloy

Alexandr Murashkin, Ming Matthew Ma
The David R. Cheriton School of Computer Science
University of Waterloo
Waterloo, ON, Canada
amurashk, m22ma@uwaterloo.ca

ABSTRACT

Sensitive data are becoming available through the Web and other distributed protocols. This heightens the need to carefully control access to data. Control means not only preventing the leakage of data but also permitting access to necessary information. Access control requires authorization rules and constraints. To express access control policies, several languages, such as XACML are used to specify which subjects can access sets of resources or services to perform specific actions.

We develop a tool based on first order logic modeling to detect and visualize possible conflicts within sets of access control policies expressed in XACML. We first translate the model into a relational first order logic language called Alloy, and then analyze interactions and conflicts among access control policies using Alloy analyzer. We then proposes potential repairs to the user thorough user interface, and automatically apply the fixes specified by the user. It is proved that with our tool can automatically determine inconsistencies in user specified model, and recommend the user and apply the repair successfully.

Categories and Subject Descriptors

D.2.4 [**Software Engineering**]: Software/Program Verification—*formal methods*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Graphical user interfaces (GUI)*

General Terms

Computer Aided Verification, Access Control Policy, User Interface

Keywords

Alloy, XACML, rule set, policy set, GUI

1. INTRODUCTION

Important data are increasingly available on-line through the Web and other distributed protocols. This heightens the need to carefully control access to data. Access control means not only preventing the leakage of data but also permitting access to necessary information. The key goal is that the right person can access right resource on the right time.

1.1 Motivation

Due to growing variety of access methods, central databases must now provide data in a large number of different contexts, each governed by specific access-control policies. Selective access control is an important mechanism in distributed system security. It can be used in order to allow a user to access only certain information, for certain purposes, at certain times, or when he or she plays a certain role. Access control is enforced by mechanisms that need to be programmed by means of policies. An organization may have many such policies, which may have been established at different times, by different people, perhaps without a clear view of all the consequences. Inconsistencies can then exist in such sets of policies.

While policy mechanisms may be able to solve inconsistencies at a higher level, users and administrators still need to be aware of them, because they may lead to unintended system behavior. For example, a policy may be added to prevent a certain access, however in fact the access is still allowed because of another policy of higher priority, or the new policy may prevent access of someone who should remain authorized. We will show in this paper that such inconsistencies can be detected and fixed by using our tools.

1.2 Background

XACML for Access Control

XACML is an OASIS standard [9] that defines an architecture, policies and messages within an access control system. XACML generally center around attributes; attributes describe subjects, actions, and resources. For example, faculty is a value for the attribute role, which describes the subject. The names of attributes, such as role are called attribute ID, while the values bound to them are called attribute values (such as Faculty). The fragment of an XACML policy designating Faculty as a role is shown below:

```
<SubjectMatch MatchId=''...:string-equal''>
  <AttributeValue DataType=''...string''>
     Faculty</AttributeValue>
   <SubjectAttributeDesignator
```

```
      AttributeId=''role''
      DataType=''...string''
   </SubjectAttributeDesignator>
</SubjectMatch>
```

A rule in XACML specifies which decision to take as a function of the attributes. The XACML takes in a request, names and values of a set of attributes, and makes an access-control decision on it based on the specified rules. A decision can be permit, deny, or not-applicable which indicates that the requests is not handled.

The access control policies are stored as XML files, they are structured into rules, policies and policy sets. Several rules are grouped into policies and policies are grouped into policy sets. Rules, policies and policy sets define a target which indicates their domain of applicability. A target specifies a set of attributes and their values that should match those given by a request. For example, if there are two policies, the first one is applied when the subject is a student, the resource is the marks file and the action is printing, and the second policy is applied when the subject is a professor, the resource is the marks file and the action is modification. If a request from a student to print the marks file comes to the policy enforcement point, the policy decision point will only select the first policy to make a decision.

Alloy

Alloy [6] is a modeling language based on relational first order logic. An Alloy model is a collection of constraints that describes a set of structure. Alloy's tool, the Alloy Analyzer, is a solver that takes the constraints of a model and finds structures that satisfy them. It can be used both to explore the model by generating sample structures, and to check properties of the model by generating counterexamples.

To model structures, Alloy uses the concepts of signature and relation. A signature is a type in Alloy. It can be considered equivalent to a class in the object oriented paradigm since a signature can be instantiated. For example, we can define a abstract general signature Element like this:

```
abstract sig Element {
         attributes : Attribute -> Value}
         {attributes in values}
sig Subject, Resource, Action extends Element{}
```

A relation is a structure that relates signatures and their instances. Functions are special binary relations; they map each instance from the left signature to only one instance from the right signature (the function effect maps each rule to only one effect). Constraints are represented in Alloy by facts. A fact is a logical formula that always holds. Alloy uses first order logic in an ASCII format. We can also specify predicates that could be evaluated to return true or false and functions that could return signature instances. Alloy is able to automatically instantiate and evaluate predicates.

1.3 Related work

Regarding access control policy verification, there are several various related approaches. [5] proposed encoding XACML access control policies ordering relations and then translation of them into SAT solver for verification. This paper defined some Alloy notation as well. The papers [4] and [8] defined XACML access control policies notation in Alloy and deal with verification problem. The latter approach - [8] - is used in our paper, it introduces the usage of predicates for access control policy model verification and validation. We found the way the models are defined in this paper straightforward, so we used this approach and extended it. Other than that, [2] describes inconsistency checking in role-based access control policies (RBAC). [3] proposes Margrave - a tool for access control policies (XACML and other formats) verification and change-impact analysis. [7] used description logics to formalize and verify access control policies. The authors also specified semi-automated access control policy repair as their future work.

None of the paper above solves the problem of access control policy repairing. It does not seem to be published prior related work in this topic. Rather than that, [11] offers access control policies verification in the language called RW and then synthesize verified specifications in XACML. [1] considers another type of access control policies - XML write access control policies - offers repairs in case of that some actions can be simulated by multiple another actions. Regarding repair in general and different domains, [10] proposes repair trees for inconsistency solving in design models. We use similar trees in our repairing approaches. There are works on model repairing in temporal logic, but we do not consider the notion of time in our work. Semi-automated repairing of XACML access control policies is the most important contribution of this paper.

2. PROBLEM STATEMENT

Although XACML has already achieved a considerable degree of industrial acceptance, by itself, it is impossible to determine inconsistencies and apply the fixes to repair the access control model. Inconsistencies exists as an organization may have many such policies, which may have been established at different times, by different people, perhaps without a clear view of all the consequences.

Determining inconsistencies require verifications and finding counter examples. This process is time consuming depending on the choice of verification and modeling language. Once inconsistencies are found, the user has to apply fixes but there is no guarantee that the fixes can indeed repair the model without introducing new inconsistencies. Thus, the entire model construction process, model verification process/ finding inconsistencies and repairing process can be overwhelming.

In this paper, we propose a tool that we developed that can automatically determine the inconsistencies, recommend the repair to the users and apply fixes to the access control model.

2.1 Paper structure

The Design overview section brief overview of our mission and solution to the problem stated; we will introduces the main components of our tool and the implementation approach at different stage. We will then present the implementation details in Implementation section where we articulate our design approach and implementations. Evaluation section contains our experiment results and analysis for the

results. In future work section, we discuss about our tool limitations and bring up possible future improvement; expert's comments are also included as a future references. Conclusion section summarizes completed work and the results.

3. IMPLEMENTATION

In this section, we present the verification and repair appoaches, and architecture of the tool.

3.1 Verification approach

The verification procedure is based on the one in the technical report [8]. We use the same idea in our paper. The first step is to define the Alloy input file.

1. First, a meta-model (or abstract model) is defined at the beginning of the file. This meta-model is domain-specific. It consists of signatures (*Policy*, *PolicySet*, *Rule*, $Effect$, *RuleCombiningAlgorithm*, etc.) and relations (*rules* relation links a policy and rules it contains).

2. Next, a concrete model is defined. It consists of signatures ($Policy1$, $Policy2$, $Rule1$, $Rule2$) and facts ($Policy1$ contains $Rule1$ and $Rule2$) regarding the relations specified in meta-model. Concrete model can vary within the given domain, so one meta-model can be used for many concrete models within the domain.

3. And finally, a property predicate is defined. This predicate is domain-specific as well, and each predicate represents the negation of the property we need to verify. In this paper, we focused on an abstract property. It is stated as follows: within a policy set that has *OnlyOneApplicable* policy combining algortihm, there is no two policies that for a given request return different decisions. The full body of the predicate is given in the Appendix, and its signature looks as follows:

```
pred InconsistentPolicySet [ps : PolicySet,
    req : Request,  p1: Policy, p2: Policy,
    r1: Rule, r2: Rule]
```

The predicate is supposed to find a policy set ps that has *OnlyOneApplicable* policy combining algorithm and contains policies $p1$ and $p2$, such that policy $p1$ contains rule $r1$, policy $p2$ contains rule $r2$. At the same time, $r1$ defines the response of policy $p1$: the rule $r1$ is either the only applicable rule in context of the given request req, or dominates all other rules of the policy $p1$ after apllying the rule combining algorithm. And similarly, $r2$ defines the response of policy $p2$.

And the second step is to send the Alloy file to Alloy Analyzer and execute the property predicate. We will denote the part of Alloy file without predicate as a model (so, the model consists of two parts: abstract model and concrete model). If the model is consistent, then the predicate is inconsistent with respect to the model, and Alloy Analyzer cannot generate an instance (example). If the predicate is consistent, it means that the model is inconsistent with respect to the property (the property is not satisfied in the model), and a counterexample is returned.

This verification approach gives certain advantages. We can localize the inconsistency, since Alloy Analyzer will show the values of all the arguments of the predicate in the generated instance, including the request req, even if there is no definition of request. It is not the case if we make the predicate denote the property itselft, not its negation: in this case, if the property is not satisifed, then $UNSAT$ cores can be shown, but it takes computation time to minimize them, and they still might not be minimal. So, for error localization the taken approch is better.

3.2 Repair

Once there is an inconsistent policy set, the verification procedure will return the policy set ps, policies $p1$, $p2$ and the corresponding rules $r1$ and $r2$. Our repair procedure is applied to policies or rules, and this will affect the consistency of the policy set ps. Since our verification procedure returns two rules of two policies, actually we can try to apply the same set of repairs to each of them by simmetry. Therefore, we can consider the following simple repair ways in the context of the returned policy set ps, and one policy $p1$ and rule $r1$ within it.

1. Switch the effect of the rule $r1$. If it was $Permit$, it is changed to $Deny$, and conversely.

2. Switch the policy combining algorithm of the policy $p1$. If it was $PermitOverrides$, it becomes $DenyOverrides$, and conversely.

3. Remove the rule $p1$ from the policy $p1$.

4. Remove the policy $p1$ from the policy set ps.

5. Switch the policy combining algorithm of the policy set ps from *OnlyOneApplicable* to either $DenyOverrides$ or $PermitOverrides$.

The ways 1 and 2 seems to be the best. Just switching the values is definitely less radical. Ways 3 and 4 are dangerous: first, they might affect other requests; next, they can bring us to the empty model which satisfies everything. Way 5 is quite global to apply it, and again, it will hide the conflicts at all, and this can cause more problems in the future. There can be other ways, though: changing the subject, action and resource, but this is risky and can increase the number of states.

So, for this project, we end up with the first two repair ways and we can apply them to both $p1$ and $p2$, so actually we have 4 repair ways:

1. Switch the effect of the rule $r1$.

2. Switch the policy combining algorithm of the policy $p1$.

3. Switch the effect of the rule $r2$.

4. Switch the policy combining algorithm of the policy $p2$.

However, the application of one repair way might not be enough, in case that the core of inconsistency is not represented by two rules only. First, there might be redundant rules. For example, if the rule $Student- > Read|Modify- > Marks$). This case is handled during the optimization during the conversion of the source file to Alloy file. Next, there are cases that a policy set is faulty for more than one request (not only for the request Professor->Read->Marks, but with the request Professor->Modify->Marks, for instance), and it definitely needs more than one repair procedure in our approach. This is why we propose the displaying of the approximate numbers of next fixes.

The approximate number of next fixes is calculated as follows. After repairs have been proposed, the system tries to apply each of them gaining partially repaired models. Then, it runs the verification again for each of the partially repaired models. If some partially repaired model is consistent, it means that this model is fully repaired. So in this case, the approximate number of next fixes is zero. If some partially model is still inconsistent, then the system tries to apply fixes to the partially repaired model again, and adds 1 to the number of next fixes. This process can run infinetely and cause state explosion problem, so we limit the depth to certain ammount (in our case we specified 2).

If the number of next fixes for each fix is greater than 0, then user needs to apply fixes multiple times. The user will go down the repair tree until he will get the fully repaired model.

In this approach, more than one path can produce the repaired model. It does not matter whether this is a minimal one or not: if the user chooses the path with minimal number of next fixes, then eventually he will get the repaired model with minimal number of changes.

However, there is a problem with this approach - cycles. For example, swithing rule combining algorithm may not be helpful. So the system will propose to switch rule combining algorithm and them, at the next step, switch it back. However, our approach helps to solve this problem: if cycles are present, then the way that leads to cycles will be annotated with bigger number of next fixes automatically, so it is less likely that the user will want to choose this way.

3.3 Overview

The function of our tool is to automatically determined inconsistencies within access control policy defined by XACML, and recommend the potential repair to the user, and finally apply the user chosen fixes automatically. Throughout this process, user only have to review the recommended fixes, and choose the fix by clicking a button in user interface, and then the model should be fixed automatically by our tool.

To increase the usability of our tool, we have an external module, a converter that can automatically generate Alloy model based on a table template (or xml, csv file template) so that the user does not need to have complete literacy in Alloy modeling language. As the content of the such a template, the user is asked to specify the rule, subject, action, resource and effect in predefined format. Our converter will than generate the Alloy model and its predicate to verify inconsistencies, which is a safetyness property. Once Alloy model is generated, we input the model into our inconsistency detection and repair tool.

3.4 Architecture and Data Flow

Our tool consists of three major parts: user interface, PropertyVerifier&Fixer and Alloy analyzer. The architecture is shown in the figure below (Fig.1) .

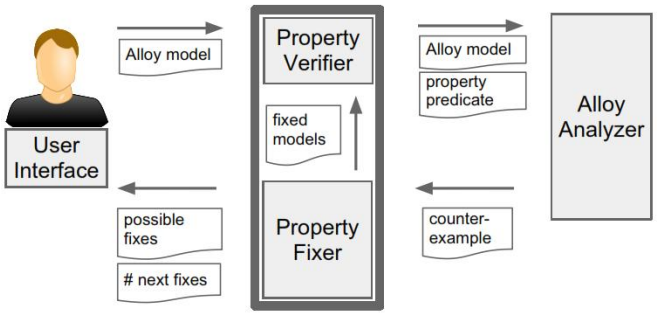

Figure 1: Tool Architecture and Data Flow

At start, the UI first takes in the Alloy model created, and sends it to PropertyVerifier&Fixer. Next, the PropertyVerifier part starts Alloy analyzer to process Alloy model and its property predicate; please note that , we write our predicate in a way that it checks for if two policies return inconsistent results. PropertyFixer, then takes in the results from Alloy analyzer and extract relevant information, such as inconsistent rules, subject, action, resource and effect, and process those information to propose the potential fixes. The fix recommendation from PropertyFixer is passed to the user interface, and the user is shown with a list of fixes. Once the user chooses one particular fix, our tool will automatically update the original Alloy model. This updated model is then passed to PropertyVerifier to check if there are any inconsistencies. This process is repeated until the model is verified to be consistent.

In order to realize the above mentioned architecture, we have created several Java modules which is summarized below.

InteractiveRepairer

This is the core of our tool, it contains:

- Main engine
- PropertyVerifier
- PropertyFixer

The Main engine, controls the communication between user interface and our Fixer and Verifier; it process input from UI and output the results to UI for user inspection. Proper-

tyVerifier calls Alloy analyzer through our API and obtain XML format of the generated instance that contains information about the policy, rule, request and effect.PropertyVerifier then follows XPATH standard and store those information into nodes for later processing in Java. Once the XPATH nodes are created with relevant information from Alloy analyzer, PropertyFixer takes in those information and proposes potential fixes. Once the user select the preferred fix, this decision is communicated to PropertyFixer again through Main engine, and PropertyFixer will automatically apply the fixes to the original Alloy model.

Note that PropertyFixer and PropertyVerifier function as a pair, and one pair is required for each property to be verified; in our study, we limit the property verification to safetyness property but this is extensible by creating another pair of PropertyFixer and Verifier.

AlloyRunner
AlloyRunner is an API that facilitate the communication between our tool with Alloy analyzer. Mainly, it initializes the Alloy analyzer, sends model to Alloy analyzer, and get XML instances from Alloy analyzer. Alloy analyzer serves as verification engine for our input model and will return instances or counter example which we use in later processing.

Presentation Layer
Presentation uses dynamic web application technologies, such as Javascripts, Node.js to allow interactive model repair. Node.js executes interactive repair with user inputs and calls back to the Main engine in InteractiveRepairer to apply the repair. For instance, the propose button triggers Main engine to start PropertyVerifier to execute the model verification, and the resolve button will cause the PropertyFixer to apply the repair on the original Alloy model. The return value associate with each function call-back of the button is error description for propose, and new model for resolve button.

During design phase, we studied all possible ways of fixing the model, and we have chosen four fixes in total considering the efficiency and relevancy of access control policy fix. From experimental results and observation, we decided to prioritize the fixes that modifies the rule effect as it reduces the number of total fixes greatly. The details are presented in the following subsection.

3.5 Model Fix
ToBeAdded

4. EVALUATION RESULT
4.1 Form and purpose
ToBeAdded

4.2 Results
ToBeAdded

5. FUTURE WORK
Although our tool demonstrate potential value in automated access control model repair, while significantly reducing user's manual work, we see future improvements in our tool. During implementation and testing stages, we noticed the following limitations in our tool:

1. Verification is bounded
2. Can verify one property in a time until first counterexample is found
3. Subset of XACML is covered
4. Repair procedure depends on the property and requires pre-defined prospective repair ways

The first two points are pretty clear; our verification is bounded. Although this enables us to use Alloy as it is for bounded model checking, we would have significant dependability to Alloy. In this case, software upgrade in Alloy would cause potential mul-functioning of our tool which needs to be addressed in the future. In this project, we mainly interfaced the communication with Alloy API therefore, in the future, we need to communicate with Alloy back-end to prevent this problem.

Although we strive to make our tool scalable, the automated Alloy model generation from a user specified table brings limitations in the structure of the model we generate. Although this access control model works well for our present study, it only represents a subset of XACML. In the future, we would like to explore different ways to increase the flexibility of our model. One potential way is to incorporate sketching techniques and having multiple rule tables so that the user can choose the structure of the model our tool generates.

State explosion is not addressed in a smart way due to time constraints. We limit the depth of the potential solution to two. This obviously will become a problem when the examples are large or more complicated. We would like to study more on ways, such as graph theory and multi-object optimization to deal better with this common problem in verification and repair.

One important future improvements is on the repair recommendation. Currently, we only proposes repair based on rule effect and override effect. We derived these to approach from observation and testing. We will explore other factors such as relation, and derivation in the future to enhance our repair framework.

6. CONCLUSION
6.1 Summary
In this report, we presented our tool which can automatically determine inconsistencies in XACML access control policy, recommend the fixes to the user and automatically apply the repair on the model. We allow user interaction by an user interface where we display our fixed results, Alloy model and also obtain user input.

We have created the input converter so that user can just specify the policies in our provided format in CSV, table form, and we will automatically generate Alloy model for the access control policy, reducing the manual work and user

knowledge on Alloy as a modeling and verification tool. Our tool takes in the Alloy model and run Alloy analyzer in the back-end, and retrieve the information on generated inconsistency instances to propose potential fixes. Among the fixes, we also show the total number of fixes required for the selected fix. From evaluation result, it is proved that our tool can help the user to repair inconsistent access control policy model effectively, requiring only user input for fix selection as a manual part.

6.2 Implications

From our project, we conclude that "automated" model repair is possible for access control policy defined by XACML structure. We make implications on the fact that user input is still necessary to reduce the redundant computation and steps taken to fix the model. By allowing user to select from list of recommended fixes, and use the user input as new argument in our automated repair tool, we are able to efficiently repair the model without looping.

One other implication we make is that our tool should be easily extensible so that the model repaired is not limited to access control policy. We see our InteractiveFixer module as a "head" which we can change according to the targeted model that needs to be verified.

7. REFERENCES

[1] Loreto Bravo, James Cheney, and Irini Fundulaki. Repairing inconsistent xml write-access control policies. In *Proceedings of the 11th international conference on Database programming languages*, DBPL'07, pages 97–111, Berlin, Heidelberg, 2007. Springer-Verlag.

[2] V. Cridlig, R. State, and O. Festor. A model for checking consistency in access control policies for network management. In *Integrated Network Management, 2007. IM '07. 10th IFIP/IEEE International Symposium on*, pages 11 –19, 21 2007-yearly 25 2007.

[3] Kathi Fisler, Shriram Krishnamurthi, Leo A. Meyerovich, and Michael Carl Tschantz. Verification and change-impact analysis of access-control policies. In *Proceedings of the 27th international conference on Software engineering*, ICSE '05, pages 196–205, New York, NY, USA, 2005. ACM.

[4] Graham Hughes and Tevfik Bultan. Automated verifcation of access control policies. Technical report, 2008.

[5] Graham Hughes and Tevfik Bultan. Automated verification of access control policies using a sat solver. *Int. J. Softw. Tools Technol. Transf.*, 10(6):503–520, October 2008.

[6] Daniel Jackson. *Software Abstractions: Logic, Language and Analysis*. The MIT Press, revised edition, 2012.

[7] Vladimir Kolovski and James Hendler. Xacml policy analysis using description logics. Technical report, 2008.

[8] Mahdi Mankai and Luigi Logrippo. Access control policies: Modeling and validation. Technical report, 2008.

[9] OASIS Open. *eXtensible Access Control Markup Language (XACML)*. The MIT Press, 2.0 edition, 2005.

[10] Alexander Reder and Alexander Egyed. Computing repair trees for resolving inconsistencies in design models. In *Proceedings of the 27th IEEE/ACM International Conference on Automated Software Engineering*, ASE 2012, pages 220–229, New York, NY, USA, 2012. ACM.

[11] Nan Zhang, Mark Ryan, and Dimitar P. Guelev. Synthesising verified access control systems in xacml. In *Proceedings of the 2004 ACM workshop on Formal methods in security engineering*, FMSE '04, pages 56–65, New York, NY, USA, 2004. ACM.

APPENDIX

A. USER EXPERIMENT TASKS AND QUESTIONNAIRE

I. Introductory presentation

II. Tasks

Tasks are split into 3 categories: **Overview**, **Graph** and **Case Studies**. All tasks are to be performed using the think-aloud protocol, that is, the participants should verbally describe what and why they are doing. All tasks are based on the file *AndroidSampleMoo_4.cfr*.

1. Overview:

T1. Upload a model file *AndroidSampleMoo_4.cfr*.
T2. Tell what features and quality attributes are specified in the Android phone model.
T3. Tell what objectives are presented in the Android phone model.
T4. Tell how many Android phone configurations are generated.

2. Graph:

T5. Identify 4 phones with the highest total performance (Just say: "P29, P30, P12, P4", for instance). What is their total mass?
T6. Identify a phone with the lowest energy consumption. What is its total performance?
T7. Identify the phones with very low total mass.
T8. How many phones have perfect (highest) total security? What features contribute to total security?

3. Cases studies:

T9. Your boss says that among all the sets of phones you need to issue only one. He says that he needs to choose one among the phones with the best total performance. He is OK with any high energy consumption, since new Android battery is very good. He says he does not need a perfect total security, but it should be more than 0. So he is OK to sacrifice total security to get more total performance.
Your actions:
1) Which product(s) will you choose and why?

T10. Your boss says that it is interesting that bubbles representing P5, P8, P2, and P22 are located close to each other. And he wants to know why this happened. Because, maybe, the phones are equivalent in some sense and we can consider them separately as an equivalence class.
Your actions:
1) What do the products have in common and what are the differences?
2) Why the bubbles are located close to each other?

T11. Your boss needs you to analyze the total performance and total mass of all prospective phones that have USB and WiFi features. He knows that P5, P8, P2, P22 and P18 have USB and WiFi. But he wants to consider all the phones that have USB and WiFi features.
Your actions:
1) Make the set to be a complete equivalence class by adding other relevant products to it.
2) Among the selected products within the equivalence class, what is the:
Minimal,
Mean,
Maximal
total performance? Are the phones differ in total mass significantly or not?
3) Based on the question above, is it reasonable for the boss sacrifice total mass to gain more total performance?

Questionnaire

Q1. Do you find the idea of representing the products in a four-dimensional space using bubble chart useful? Any comments on that? Are there limitations? Does the representation of the four dimensions make it easy to find "better" products?
Q2. What do you think about the notion of the complete equivalence class? What is it and how it can be used? Describe it in your own words.
Q3. I noticed that I can group products together based on their features, so each cluster will be denoted by an equivalence class. Let's call it "clustering by features". Do you think it is a good idea or not?