Master in Artificial Intelligence

Master of Science Thesis

Norm Crowdsourcing

Author:
David Sánchez Pinsach

Supervisor:
Dra. Maite Lopez (MAiA UB)

Cosupervisor:
Dr. Juan Antonio Rodríguez (IIIA-CSIC)

Facultat d'Informàtica de Barcelona (FIB)
Facultat de Matemàtiques (UB)
Escola Tècnica Superior d'Enginyeria (URV)

Universitat Politècnica de Catalunya (UPC)
Universitat de Barcelona (UB)
Universitat Rovira i Virgili (URV)

Defense date: April 30, 2015

April 2015

"As we look ahead into the next century, leaders will be those who empower others."

– Bill Gates

Abstract

Empowering users to participate in processes that up to date are controlled by others represents a highly interesting research topic. Nowadays, a vast number of users are connected to social platforms in order to share different types of contents such as images, posts or comments. However, the actual regulation, management and content control of these systems are based on the figure of moderators and are not working correctly, due to the fact that it is required to have a large number of them and their work is time-consuming. Thus, it is necessary to introduce users in the process of moderation in order to reduce moderators and to guarantee that users themselves decide what is allowed or not in their social community.

The aim of this Master Thesis is to define a method to support users with the co-creation of their own norms in those virtual social communities where they belong. In this sense, a dedicated norm argumentation method is proposed and used to structure and facilitate users' interaction as well as to aggregate different opinions about positive and negative arguments that will support the activation of new norms in the community.

Finally, the system is tested with human users in order to view the feasibility of the proposed solution.

Resum (Catalan)

Empoderar els usuaris per participar en processos que fins ara estan estaven controlares per altres representa un tema de recerca de gran interès. Avui dia, un gran nombre d'usuaris estan connectats a plataformes socials i comparteixen diferents tipus de continguts com imatges i comentaris. Malgrat això, els sistemes actuals de regulació, administració i control del contingut basats en la figura dels moderadors no funcionen correctament, donat que se'n necessiten molts i el seu treball requereix temps. Així doncs, és necessari introduir els usuaris en aquest procés de moderació per tal de reduir els moderadors i perquè ells decideixin que esta permès o no en la seva comunitat virtual.

L'objectiu principal d'aquesta Tesi de Màster és definir un mètode per ajudar als usuaris amb la creació de les seves pròpies normes en la comunitat social virtual a la que què ells pertanyen. En aquest sentit, un mètode dedicat d'argumentació ha estat proposat i utilitzat per estructurar, facilitar la interacció dels usuaris i per agregar diferents opinions sobre arguments positius i negatius que donin suport a la creació de noves normes a la comunitat.

Finalment, el sistema ha estat provat amb usuaris reals per tal de veure la factibilitat de la solució proposada.

Resumen (Spanish)

Empoderar a los usuarios para participar en procesos que hasta ahora estaban controladas por otros es uno de los temas de investigación de gran interés. Hoy en día, un gran número de usuarios están conectados a plataformas sociales y comparten diferentes tipos de contenidos como imágenes y comentarios. Sin embargo, los sistemas actuales de regulación, administración y control del contenido basados en la figura del moderador no funcionan correctamente, dado que se necesitan muchos y su trabajo requiere tiempo. Así pues, es necesario introducir a los usuarios en este proceso de moderació para reducir los numero de moderadores y para que ellos decidan que esta permitido o no en su comunidad virtual.

El objetivo principal de esta Tesis de Máster es definir un método para ayudar a los usuarios con la creación de sus propias normas en la comunidad social virtual a la que ellos permanecen. En este sentido, un método dedicado a la argumentación ha estado propuesto y utilizado para estructurar y facilitar la interacción de los usuarios además de agregar diferentes opiniones sobre argumentos positivos y negativos que den soporte a la creación de nuevas normas a la comunidad.

Finalmente, el sistema ha estado provado con usuarios reales para ver la factibilidad de la solución propuesta.

Acknowledgements

I would like to thank my supervisors, Dra. Maite Lopez Sanchez and Dr. Juan Antonio Rodriguez Aguilar to give me the opportunity to work with them. They bring me a huge support, perfect guidance, strong encouragement to any moment and I think that this work could be not possible without them.

Moreover, I would like to thank my family and friends for the support in the difficult moments, specially to my mum Teresa and my girlfriend Csilla.

Finally, I want to thank the Self-Organisation Agent System students and also Javier Morales, Iosu Mendizabal and Maria Salomo for the time that they spend in my tests and their comments about the system.

List of Figures

2.1 Appgree webpage 10
2.2 Liquidfeedback homepage 11
2.3 Stackoverflow webpage 12

3.1 Structure of the system 14
3.2 Structure of the norm 15
3.3 Flow of the website 17

4.1 User case 22
4.2 Forum section 23
4.3 Reporter section 24
4.4 Image/Videao section 25
4.5 Upload content 26
4.6 Alert triggered the user tries to upload prohibited 26
4.7 Active norm 27
4.8 Dababase table structure 27

5.1 Users of the test 30
5.2 Experiments distribution 32
5.3 Norm 1 from first experiment 35
5.4 Norm 2 from first experiment 37
5.5 Norm 3 from first experiment 37
5.6 Norm 4 from first experiment 38
5.7 Norm 1 from second experiment 38
5.8 Norm 2 from second experiment 39
5.9 Norm 3 from second experiment 39
5.10 Questions 40
5.11 Box plot 41

List of Tables

5.1 General traits of human testerss 31
5.2 Questionnaires' results 36
5.3 Statistical results 36

Contents

1 Introduction **1**
- 1.1 Motivation . . . 3
- 1.2 Objectives . . . 4
- 1.3 Associated publications to this Master Thesis . . . 5
- 1.4 Master thesis structure . . . 5

2 State of the art **7**
- 2.1 Research Literature . . . 7
  - 2.1.1 Norms . . . 7
  - 2.1.2 Simulation systems . . . 8
  - 2.1.3 Argumentation . . . 8
  - 2.1.4 Aggregation . . . 9
- 2.2 Web sites and applications . . . 9

3 Methodology **13**
- 3.1 Conflict . . . 13
- 3.2 Structure of the system . . . 14
- 3.3 Norm life cycle . . . 15
  - 3.3.1 State diagram . . . 16
  - 3.3.2 User participation . . . 17
- 3.4 Method to aggregate user opinions in norms . . . 18
  - 3.4.1 Norm argument maps . . . 18
  - 3.4.2 Voting aggregation . . . 18
  - 3.4.3 Semantic . . . 20

4 Development **21**
- 4.1 Functionalities . . . 21
  - 4.1.1 Sections, Users, Contents, Complaints . . . 21

4.1.2 Technologies 22

5 Evaluation **28**
5.1 Configuration 29
5.2 Norm activation criteria 32
5.3 Experiments 34
5.3.1 Experiment 1 34
5.3.2 Experiment 2 34
5.3.3 Results 35

6 Conclusions and Future work **44**
6.1 Future work 45

Appendices **50**

Chapter 1

Introduction

Nowadays, a vast number of users are connected by social platforms and share different types of contents such as images, posts or comments. These social platforms store a large amount of contents whose upload requires some management or regulation. As a consequence, some monitoring is required so to follow the policies or rules of the social platform.

However, the classical regulation, management and control of the content is based on the figure of moderators and suffer from several drawbacks that difficult the task of both moderators and regular users in a social community. One the one hand, the time is limited and a single moderator can only handle a few hundreds of short contents. So, if the community grows up and it generates a lot of content more moderators will be required, increasing the cost. Additionally, a moderators' structure is required in order to supervise and coordinate the task of moderators, so that all moderators follow the same rules and apply the same punishments. On the other hand, users of a social community demand moderators act fairly, leading by example, showing respect to all users and solving those problems that may arise in the community.

Currently, a significant number of users start campaigns to change those rules in social communities that seem not to be fair enough. Therefore, it becomes necessary that users gain control of their own social communities, deciding their own rules. Nevertheless, how users can decide their own social norms in a virtual social community remains an open question in the research community.

A lot of researchers are working on trying to answer this question, providing tools, algorithms and theories to guarantee that users can decide their own norms in these environments. In this sense, this Master Thesis is in fact a continuation of a previous work on the simulation of virtual communities [12] which includes a method for automatising the generation of norms [11]. The agents in this simulation are capable to make their own decisions and actions. They can upload, view and complain about content in three sections of the virtual web (forum, reporter and image/video). The content is pre-defined and pre-labeled in order to know which type of content the agents are uploading. The system considers the following content type labels: spam, ok, violent, insult, wrong placement, and porn. Furthermore, this simulator is endowed with an advanced setup/configuration interface which allows to customise experiments by defining the user population with different number of users, different types of users, different behaviours and so on. Additionally, this virtual community simulator was endowed with an automatic intelligent system named IRON [9] that it is capable to monitor the interaction of the agents in the virtual community environment as well as to detect those conflicts (i.e., undesired situations) that may arise between agents. When IRON detects a conflict it generates a norm in order to void this conflict in the future. Also, IRON is capable to measure if the norm is necessary and useful. However, this automatic approach to norm generation is performed by an intelligent system and fails at involving users explicitly in the process of deciding what norms should regulate the community they belong to. Therefore, it is necessary to empower users in the creation of their own norms in their virtual social communities.

In this sense, this thesis tries to give a new methodology, tools, samples and theory on how to implement a solution to this moderation problem.

1.1 Motivation

Imagine a world where people can actually give their opinion about the norms and rules and these opinions have the same weight for all users. Imagine a community that ask users about its norms without imposed intermediate people. Imagine that the aggregation of the users' opinions (in the form of arguments) leads to the creation of new norms.

This thesis is aimed at establishing the basis for this to be possible, so that the current conception of democracy becomes transformed. Furthermore, this transformation could even be transferred from virtual communities to the streets in our real world. Unfortunately, current systems do not work properly and are susceptible to be corrupted by specific users such as moderators or administrators. This thesis aims at empowering community users, the ones actually using the virtual social community they belong to. The actual system do not work properly and normally is corrupted by a few of users like moderators or admins. We need to give the power to the users, to the community who are the people that are using the virtual social community.

Users have problems where they try to understand and express their opinions about the regulation of social webs because these regulations normally are showed as a long technical text that just a few users read and still less understand correctly. Furthermore, a lot of social webs have a simple complain button usually labelled as "denounce" but it is not enough to capture the needs of the users which want express more. This master thesis aims at overcoming previously mentioned limitations by proposing a method where actual users propose and discus regulatory norms in the social communities they belong to.

1.2 Objectives

This master thesis aims at defining a method to support users with the co-creation of their own norms in the virtual social communities they belong to. For this co-creation to be possible, norms have to be discussed among different users, so that a dedicated argumentation method should be properly designed. Additionally, the norms and arguments have to be structured in order to facilitate the users' interactions. Thus, to summarise, the main goals of the project are the following:

- To define a method to support users with the co-creation of their own norms.
- To built a system in which the real users can co-create their own norms.
- To propose a dedicated norm argumentation method.
- To aggregate different a opinions about positive and negative arguments by means of suitable aggregation operator.

1.3 Associated publications to this Master Thesis

This Master Thesis is a project developed in collaboration of the UB and IIIA-CSIC with several experts and researches like Maite Lopez, Juan Antonio Rodriguez, Javier Morales, Iosu Mendizabal and me. AI have got the opportunity to participate in a number of publications. The following list shows the results of this collaboration:

- Moderación Automática en comunidades virtuales (Final degree project) [12]
- Applying IRON to a Virtual Community Scenario (Article at CCIA proceedings) [9]
- Using IRON to build frictinless online-communities (Article at AAMAS'15 proceedings) [11]
- Extending NormLab to spur research on norm synthesis (Article at AI Communications journal)) [10]

In addition, another article is under elaboration for next CCIA.

1.4 Master thesis structure

This document is structured following the standard requirements of this Master Thesis. Initially, the introduction chapter situates the context of the work, and introduces both the motivation and objectives pursued by this research work. The second chapter is related to the state of art, where the related work is situated and briefly described. Related work is presented from two different perspectives. Firstly, it is related with state of the art literature, reviewing papers and research works where authors introduce related problems and propose some solutions. Secondly, this related work chapter covers actual commercial web sites in order to overview the perspective (and proposed solutions) taken by current sites. Next chapters, third and forth, are more technical and explain the methodology and implementation of the solution. Chapter 5 is related to the evaluation, where the test methodology is explained and implemented. Experiments are showed and the results are

discussed in detail. Finally, the document gives the final conclusions and future work. In particular, the conclusions recall the overall performed work and links it with the objectives from the introduction so to assess if the main objectives have been successfully achieved.

Chapter 2

State of the art

Several contributions related to this Master Thesis can be found in the market and in the literature. Co-creation of norms by the users in a social community is a novel field that is related with norms, deliberative systems, user empowerment and information aggregation methods. This chapter will present a review of the available work in this field from the research literature and the actual applications or webs.

2.1 Research Literature

A number of research works deal with parts of the research issue covered by this thesis. Subsequent subsections review norms, simulation systems, argumentation and aggregation.

2.1.1 Norms

Norms belong to a long established research area that originated from philosophers and sociologists. It is one of the most interdisciplinary fields because it is associated to behave in a specific way. In our everyday lives, we use a range of techniques for coordinating activities. One of the most important is the use of norms and social laws. A norm is an established, expected pattern of behaviour like form a queue when waiting for a bus that is not enforced norm. In addition, social laws usually carry with them some authority [14].

Recent works are focusing on the creation of the norms in social communities context. For instance, in the document [5] the author reviews and

distinguishes social from offline to online. In this sense, the author reflects that they are very similar but some points like the physical appearance can influence the social norms in the offline that in the online does not exist. In addition, the author looks into actual websites to see how the socials norms are applying, and his conclusions are twofold. Firstly, the social norms evolve as quick as the technology and thus it becomes necessary a quick adaptation of the websites to the social changes. Secondly, the author considers that some websites like Stack Overflow [1], Reddit [2] or Hacker news [3] use voting as a regulatory mechanism but there are disadvantages in how these systems work because there is a lack of homogeneity on how to perform this moderation task. Furthermore, as the author says *we need systems that preserve real world mappings of behaviour*.

2.1.2 Simulation systems

In [9],[11] the authors propose a methodology and simulation framework in order to build automatically the norms in a social community by an intelligent system who is capable to detect conflicts. Once the system detects a conflict, it tries to create a norm in order to avoid this conflict in the future. In addition, the system is capable to generalise the norms which reduces the number of the resulting norms.

2.1.3 Argumentation

Managing large-scale deliberation is a novel field. In the paper [7], Mark Klein proposed to use attention-mediation metrics in order to guide the focus of the users to allocate their efforts, in the large-scale argumentation context. He proposed several metrics to measure where users can contribute more based on their historical actions and preferences. Moreover, the author proposed the argument maps which are a way to structure the discussion focusing on arguments. An argument map has a tree-like structure where the question is on the top.

[1] Stack Overflow: http://stackoverflow.com
[2] Reddit: http://www.reddit.com
[3] Hacker news: https://news.ycombinator.com

2.1.4 Aggregation

Ordered Weighted Aggregation (OWA) operator is wide studied field within the decision making research area that was started by R.Yagger [15] in 1988. It is an aggregation operator in which several information is represented by a single value. It works basically with two vectors, where the first is the information that is needed to aggregate and the second is a weight vector. Thus, each weight element in the weight vector is assigned to each value of the information vector. One of the important issues in the OWA operator that have been studied is to determine the associated weight vector. A lot of literature exists that studies how to compute this vector [4], [2] [13].For instance, Yager [16] proposed a simple way by using the idea of the fuzzy linguistic quantifier. The author proposed a function to generate the weight that is related to the position of the information vector. Basically, weights are determined by the position, although other parameters are also considered to compute the final outcome. In the paper [3], the authors proposed to determine the weights by using linear functions. In order to compute the linear functions, the authors propose the following function:

$$f(x) = a \times x + \frac{1}{n} - a(\frac{1+n}{2})$$

where a is the interval of the function and x the value to aggregate.

2.2 Web sites and applications

Besides research works, the actual web is populated with a number of real applications. This section introduces a representative selection.

- **Appgree:** Appgree is a web and mobile application that allows groups of people to post a question and receive the most voted answer. Figure 2.1 shows a screenshot of the web interface of this application. In this system, a user can make a question and other users can provide answers to this question. The voting works in phases in which users can only do some specific actions. Firstly, the system starts with the first phase that consists on providing answers to the question. Secondly, the system requests in the second phase the rating of the answers. Finally, it is computed all the rating and the question resume winner best three options. Although the structure of phases could be viewed

as a structure process of the voting, it has several drawbacks. One of the problems is related to the time. Time determines when a phase is finished, so the users need to be aware that they have a limited period of time to perform their actions. Secondly, the system is very restrictive because the actions at each phase are limited and it is impossible to do actions such as adding answers when the system is in the answer-voting phase. Both issues mentioned before (i.e., time and limited action on the time) do not guarantee a fair voting since fixing when users are allowed to take decisions seems not to be entirely democratic because people do not take decisions when they want.

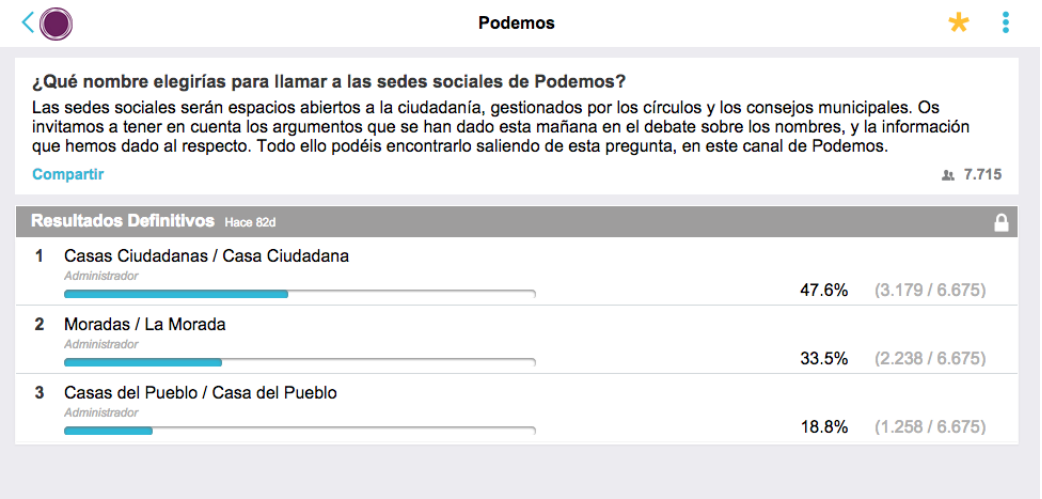


Figure 2.1: Appgree webpage

- **Liquid feedback** Liquid feedback is an open-source software that implements the concept of liquid democracy whose potential consumers are governments, corporations, political parties and associations.

  Liquid democracy (see Figure 2.2) follows the principle that assumes people do not have enough time to exercise their vote and it is necessary to delegate to someone else the responsibility of their vote. The delegated user (i.e., agent) then decides and votes by himself as well as for the other delegating users. In addition, the transfer of responsibilities can be in parts, people can determinate that for some topics someone is better to manage their vote than others, so it is possible to have several agents depending on the topic of the voting. Although the authors defend their system is fully transparent to decision process, their system has several drawbacks. On the one hand, it is difficult to demostrate that the person who is in charge of deciding by others would take the same decisions if they made their own votes, because two users may share their basic ideas but vote differently. On the other hand, it is difficult to assign responsibilities to the votes because it remains unclear who is responsible of the votes, the person who is the spokesman or the owner of the vote.

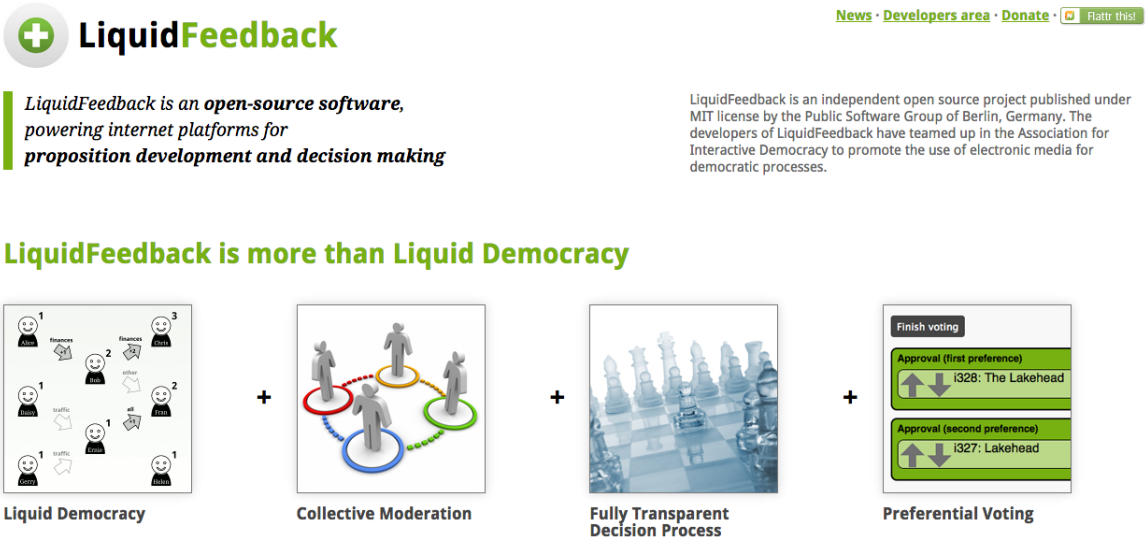


Figure 2.2: Liquidfeedback homepage

- **Stackoverflow:** Stackoverflow (see Figure 2.3) is a free question-answer (QA) site for professional and amateur programmers. The main focus is to provide answers in form of scripts, pieces of code, snippets and so on to help the programmers' issues. One of the most valuable characteristics of this platform is based on the rating of the answers and the possibility that the users who have generated the question mark which answers help him/her to solve the question. This possibility is important when you are searching for a solution to a common problem and you have a lot of answers and possible solutions. However, in this case is difficult to manage so much information and thus, you will need to scroll down a lot in the search of the proper solutions. Other important problem of this page is related on how the web shows the information. The structure does not help to navigate within the page and the discussion is difficult to follow because the system includes the question, answers and comments to the answers.

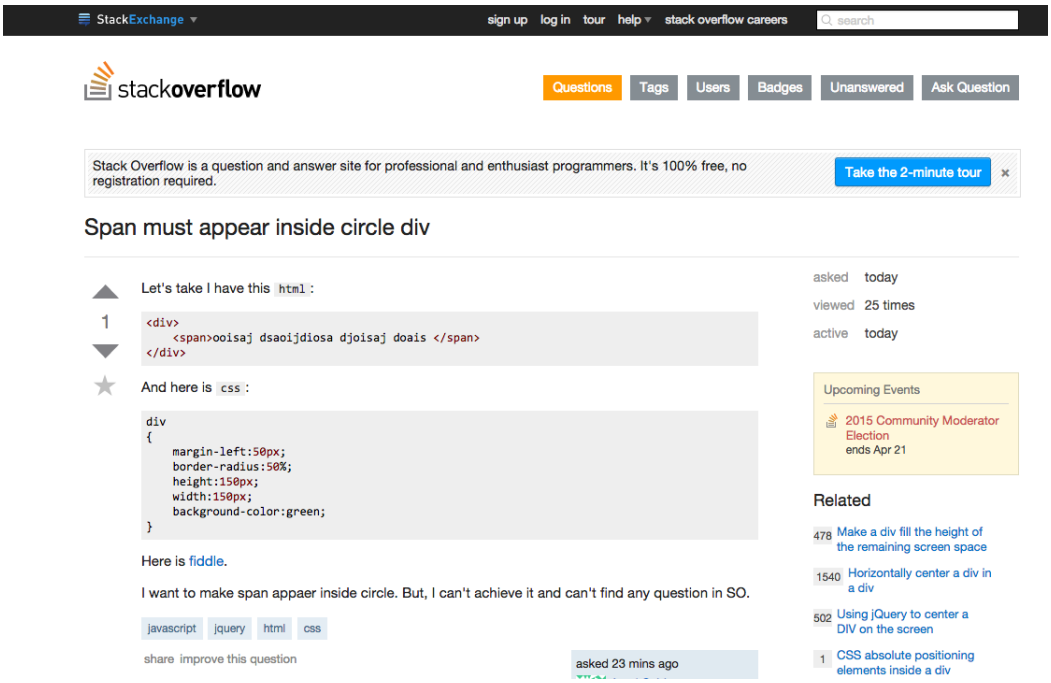


Figure 2.3: Stackoverflow webpage

Chapter 3

Methodology

This section will describe the proposed solutions to face the lack of the user participation in the process of the creation of their own norms in their social community. Also, it is reviewed different states of content and norm.

3.1 Conflict

The overall problem of moderation in a virtual social community is based on the concept of conflict. This work considers a conflict as an undesirable situation that rises in the virtual social community. Conflicts thus become the main problem that virtual communities face when users perform actions that are not considered appropriate such as, for example, upload inappropriate content, such as spam, violent or pornographic content.

Nowadays, conflicts are managed by moderators without user participation. Moderators try to identify which pre-defined norm in their community was not complied when the uploading of a content generated the conflict. Then, they apply punishments to the contents or to the users. The lack of user participation in this process is one of the actual topics that researchers are trying to solve. A novel solution/methodology to promote the user participation in the process of deciding their own norms will be defined in the following sections.

3.2 Structure of the system

Figure 3.1 shows the structure of the whole proposed system. The content showed on the top of the image has a title, the owner and two indicators that show the number of complaints and , in case a conflict is detected, what norm regulates this content. The second layer is related to the conflict. If the content has a conflict, then it becomes necessary to regulate such conflict by means of a norm. One conflict can be solved using different norms although it is preferable to use the norm receiving more support from the users. So, norms are located in the third layer. Notice however that although the system may contain several norms, a specific conflict type in our system is regulated by a single norm. Finally, the bottom layer is composed by norm arguments provided by users. The system joins norm arguments in two types: positive (pros) and negative (cons).

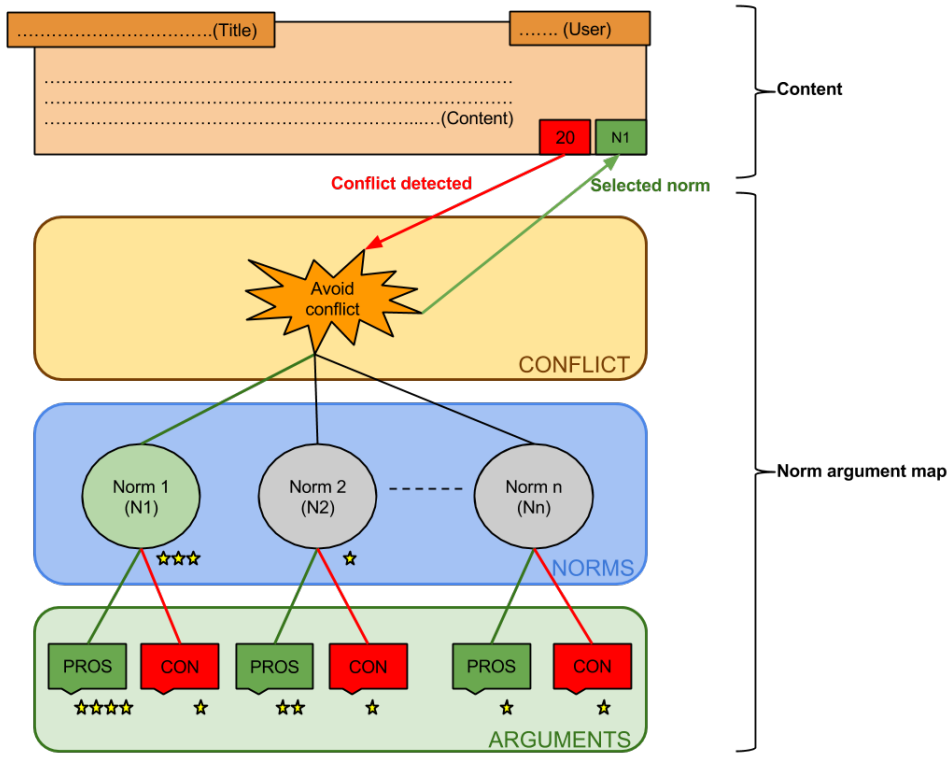


Figure 3.1: Structure of the system

3.3 Norm life cycle

Figure 3.2 shows the structure of a norm. Usually, virtual community web sites have different sections meant to include different types of content. Therefore, norms for such sites regulate the content types hat users can upload at each specific section.

Norms are of the form $\langle(\varphi, theta(\alpha))\rangle$ where φ is the norm's precondition, θ is a deontic operator (e.g., a prohibition) and α is an action that can be performed by agents (belongs to the set Ac of all possible actions). The precondition of a norm is a set of first-order n-ary predicates $p(\tau_1, ldots, \tau_n)$, where p is a predicate symbol and $\tau_1, ldots, \tau_n$ is a set of terms. The norm is defined in terms of predicates $\text{pn}(\tau_1, ldots, \tau_n)$ that in our case happen to be unary $p(\tau)$. In our case we consider two predicates $p \in section, contentType$, specifically: $p_1 = section()$ and $p_2 = contentType()$. Corresponding terms τ for these predicates are, respectively: section names $\tau_1 \in$Forum, Multimedia, The Reporter; and content categories $\tau_2 \in correct, spam, troll, rude, violent.$ and content category (τ=correct, spam, troll, rude, violent). On the other hand, a norm consequence specifies the prohibition (in this case, deontic operator θ = phr) where phr could be Prohibition, Permission and Obligation, although the system only works with Prohibition to perform the action upload of the content in the context. Therefore, norms establish prohibitions to upload certain types of contents in some sections of the community. Norms n1 and n2 are examples of normsin our scenario:

$n1 :< section(Multimedia), contentType(spam)), prh(upload(content)) >$
$n2 :< section(Forum), contentType(violent)), prh(upload(content)) >$

Norm n1 prohibits to upload spam contents on the Multimedia section, and norm n2 prohibits to upload violent contents on the Forum section.

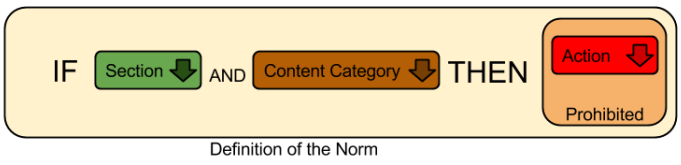


Figure 3.2: Structure of the norm

3.3.1 State diagram

Figure 3.3 shows the state diagram in terms of uploaded content and norms. On the top of the image states are associated with the content and the bottom part of the figure is associated with the norm regulating it. Firstly, when the content is uploaded, the users can view or complain about it. When several users complain about a single content, it generates a conflict that requires regulation. The regulation starts with the creation of a norm and finishes when the norm is cancelled or activated. Between these two steps, users can join a norm discussion. The norms are associated to the content that generated the conflict because the main objective is to regularise the situation of this content. The user within the discussion of the norm can perform two types of actions: add or vote positive and negative arguments. A positive argument is an argument that tries to support the necessity of a norm and the negative one tries to do it the other way around.

More specifically, when a content is uploaded the user only can view or complain about it. A threshold determines when the number of complaints related to the number of views is large enough and then is when the conflict starts. The content is arguable and marked as conflicting. At this moment, users cannot complain about content because the conflict is already detected and, thus, they only can start the discussion by creating a norm. When a user creates the norm, the content state is transformed into "under discussion". The discussion starts with a norm together with a positive argument associated to it. This associated argument of the norms justifies the norm existence with the very same content that triggered its creation. Thus, upon the creation of the norm, the system creates a positive (pro) argument that references the content that the norm tries to regulate. Subsequently, the other users can join the discussion by doing two types of actions. They can express their opinions rating other users' arguments or by adding new arguments to the discussion. New arguments can be added within in a positive or negative group. After some iterations, another threshold decides when the discussion finishes. Then, the norm could be active or cancelled and the active norms are viewed and regulated. An active norm regulates both new contents that users will upload in the future as well as those contents already uploaded in the community (which, in turn, will become marked as "prohibited").

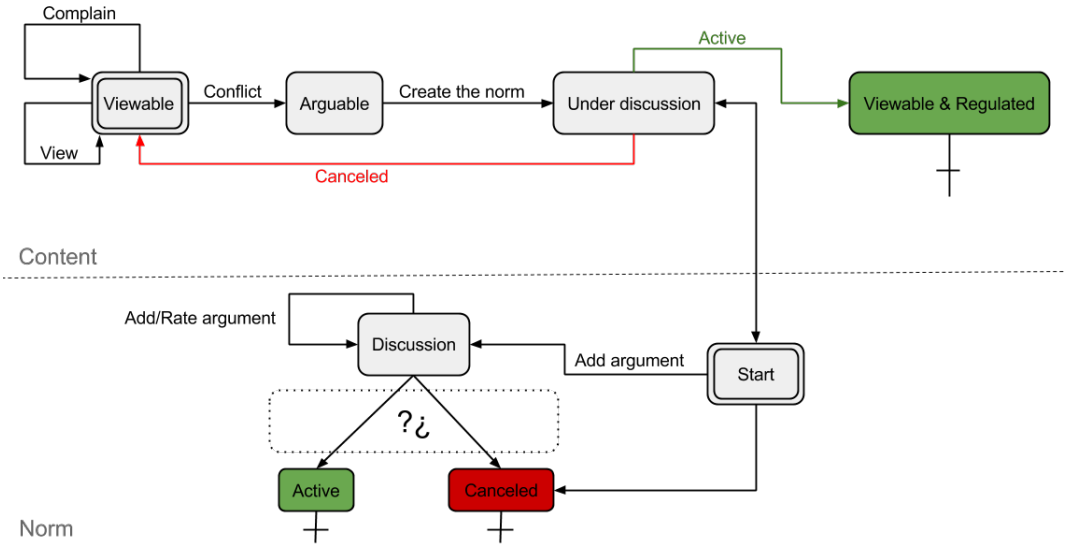


Figure 3.3: Flow of the website

3.3.2 User participation

As the Figure 3.3 shows, users participate in the creation of the norms from the very beginning. The participation starts when a user complains about a content. This is a very important point since, as previously mentioned, norms are meant to avoid conflicts (i.e. undesirable situations), and thus their existence is solely based on the uploading of conflictive contents (i.e., those that accumulate enough complaints from users in the community). So, if the system does not have conflicts, the norms will not be necessary. Once the norm discussion starts, users can participate giving their arguments to defend or reject the necessity of the norm. As a normal democracy, users can add several arguments to the norm but each argument can only be voted once by each user. If a user could vote arguments several times, it would be an unfair system due to its vulnerability to being manipulated. However, users can have several arguments to defend a norm and restricting this point may mean limiting their participation.Thus, the system allows each user to add several arguments.

The possibility to add several arguments to a norm has some implications. For instance, an unfair user could add several positive arguments to defend

the necessity of a norm and this may count as having placed several votes. As next section details, the proposed solution to this possible problem consists on adding an extra parameters to the activation of the norm.

3.4 Method to aggregate user opinions in norms

One of the most difficult challenges in the creation of norms is on how to build a norm structure to facilitate the interaction with and the interpretation of the norm. This thesis aims at tackling this challenge by proposing a refinement of Argument Maps concept by Mark Klein.

3.4.1 Norm argument maps

Mark Klein proposes the Argument Maps [7], to discuss questions in a (QA) system. An argument map is a tree-like structure where the leafs are the arguments and the root is the question. Our proposal is thus to apply Argument maps to a norm argumentation context because the discussion in this type of structure are followed easily and occupy the minimum space. Additionally, following the focus of the discussion from the very beginning, when users see the norms, is one of the most valuable characteristics of the system.

We define Norm Argument Maps as the concept of argument maps applied to norms. Bottom part of Figure 3.1 3.1 depicts its structure, which is composed by a root element that corresponds to the norm under discussion. Then, the arguments are separated into two groups (positive and negative). Each argument that is defined within this two groups will be considered as having the type of the group. So, if the user adds an argument in the positive group (pros) it will be considered as an argument that tries to defend the necessity of the norm. And analogously, so, if the user adds an argument in the negative (cons) group t will be considered as an argument that tries to diminish the necessity of the norm under discussion.

3.4.2 Voting aggregation

Each norm under discussion could have several arguments and each argument, several votes. The system needs to deal with the votes and it is a difficult task. For instance, one argument may have 100 votes which, when

computing the mean will result in a value of 4. Another argument may have 5 votes but with resulting mean of value 5. What is the most important argument, the argument with larger number of votes or the one that has higher votes? This type of questions related to how the system should aggregate the information becomes crucial to capture correctly people's opinion. Hence, we propose to use the OWA operators in order to aggregate different votes of users in the discussion of norms. The OWA operator provides a low computational method to aggregate information. Specifically, considering the group of norms N, we propose to compute the total rate of a norm $n \in N$, $totalRate(n)$, by applying the following formuli:

$$
\begin{aligned}
&A(n) = \langle A^n, \bar{A}^n \rangle \\
&\quad A^n = (a_1^n, ..., a_p^n) \quad \forall a_i^n \in A^n, i = 1..p \quad a_i^n = (us_1^n, ..., us_r^n) \quad median(a_i^n) = m_i^n \\
&\quad \bar{A}^n = (\bar{a}_1^n, ..., \bar{a}_q^n) \quad \forall \bar{a}_i^n \in \bar{A}^n, i = 1..q \quad \bar{a}_i^n = (\bar{us}_1^n, ..., \bar{us}_s^n) \quad median(\bar{a}_i^n) = \bar{m}_i^n \\
&\quad \mu^n = (m_1^n, ..., m_p^n) \\
&\quad \bar{\mu}^n = (\bar{m}_1^n, ..., \bar{m}_q^n)
\end{aligned}
$$

$$OWA(\mu^n) = \sum_{j=1}^{p} w_j \cdot b_j \qquad \text{where } b_j \text{ is the } j^{th} \text{ largest value of } \mu^n$$

$$OWA(\bar{\mu}^n) = \sum_{k=1}^{q} \bar{w}_k \cdot b_k \qquad \text{where } b_k \text{ is the } k^{th} \text{ largest value of } \bar{\mu}^n$$

$$totalRate(n) = OWA(\mu^n) - OWA(\bar{\mu}^n)$$

$$Active(n) = \begin{cases} true & \text{if } totalRate(n) \geq \beta \\ false & \text{otherwise} \end{cases}$$

Each norm n is associated to a set of positive and negative arguments A^n and $\bar{A}^n$. Previous formuli assume there are p and q positive and negative arguments respectively. Additionally, each argument a_i^n has associated a median (m_i^n) of the votes given to it by all the users (votes for an argument

i are denoted as $(us_1^n, ..., us_r^n)$). Analogously, the same considerations apply for negative arguments and their associated medians ($\bar{a}_i^n$ and $\bar{m}_i^n$). Users' votes are then used to define the argument information to aggregate. In this manner, vectors μ^n and $\bar{\mu}^n$ are defined as argument information vectors.

As previously introduced in Section 2.1.4, the OWA operator is an ordered combination function of the argument information weighted by specific weights. In our case, argument information corresponds to medians of the argument votes, but the weight vector w is also computed considering these values. Specifically, weights are calculated using the following linguistic quantifier method [16] where a and b are used to establish the limits of the function:

$$w_i = Q(i/n) - Q((i-n)/n), i = 1, ..., p$$

$$Q(x) = \begin{cases} 0, & \text{if } x < a \\ (x-a)/(b-a) & \text{if } a <= x <= b \\ 1, & \text{if } x > b \end{cases}$$

The result of the norm ($totalRate(n)$, i.e., its final rating) is computed by applying the OWA operator to both the positive and negative argument information vectors and substracting the one for the negative arguments from the positive one.

A priori, a difference between positive and negative arguments higher than β would be enough to determine that the norm positive arguments surpass the votes of the negative ones, and thus, to decide that the norm should become active. Nevertheless, this criteria is weak in some scenarios where users try to manipulate the votes. For this reason a minimum number of arguments provided by different users will be also required. Section 5.2 details the specific condition that has been actually required in the system.

3.4.3 Semantic

When computing the weight vector, parameters a and b serve to cut the extremes of the function. As a consequence, outlayer votes can be avoided. The chosen computation puts the emphasis to the order of voted values. Thus, the higher the vote is, the more weight it gets. This is so because in this norm argumentation process we give more importance to strong arguments, no matter if they are in favour (pro) or against (cons) a norm. This has the advantage of helping to avoid people that tries to manipulate the system.

Chapter 4

Development

This chapter is devoted to briefly introduce the development of the proposed methods, framework, databases as well as the look and feel of the different parts of the social website.

4.1 Functionalities

The users can perform several types of the actions within the website. This actions are showed in the user case 4.1.

4.1.1 Sections, Users, Contents, Complaints

Despite the fact that all contents in the platform are conceptually the same, each section in the platform shows them in different ways, which depend on the type of content and section. The reporter section looks like a blog where the content is showed in boxes 4.3. However, the forum section looks different. The content is situated in topics where people discuss different topics 4.2. The last section, is called image/video where the media content appears 4.4.

In order to facilitate the creation of the norms, the content is predefined and labeled. Each content has an specific type that is used to create the norms. So, if the user aims at uploading content, he/she will choose a specific button (the one corresponding to the type of content that he/she is willing to upload) and then, the system shows a list of contents of this type

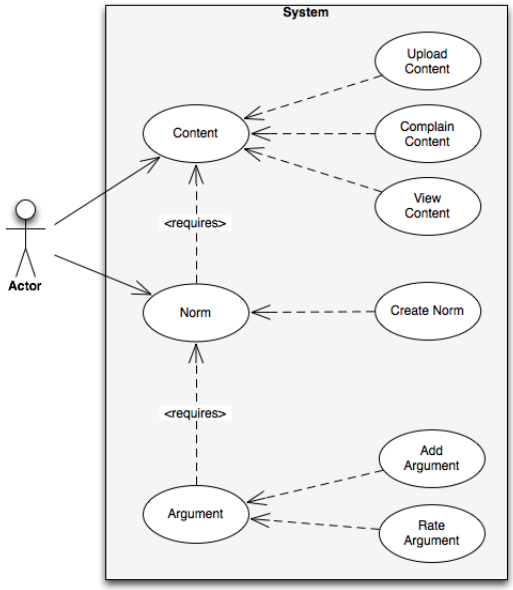


Figure 4.1: User case

(see Figure 4.5) from whom the user can select the one to actually upload. 4.5.

Finally, when a norm is active (see, for example, the norm in Figure 4.7), then the content of a specific type and located at the given section is regulated. Consequently, if a user located at a specific section, uploads a content whose type is prohibited for this section, then the system automatically pops up an alert (see Figure 4.6) informing the user that he is about to infringe a norm. The user is then free to cancel the uploading so to comply with the norm or to continue with the upload (thus infringing the norm). In the former case, the user will be congratulated to comply with norms, whereas in the latter case the content will be uploaded but labeled as prohibited. 4.6

4.1.2 Technologies

The deployed system has been designed based on a client-server architecture so to guarantee ease of access. Thus the application is hosted at a web server and users access it by means of their regular web browser. This application has been built by using the following tools and technologies:

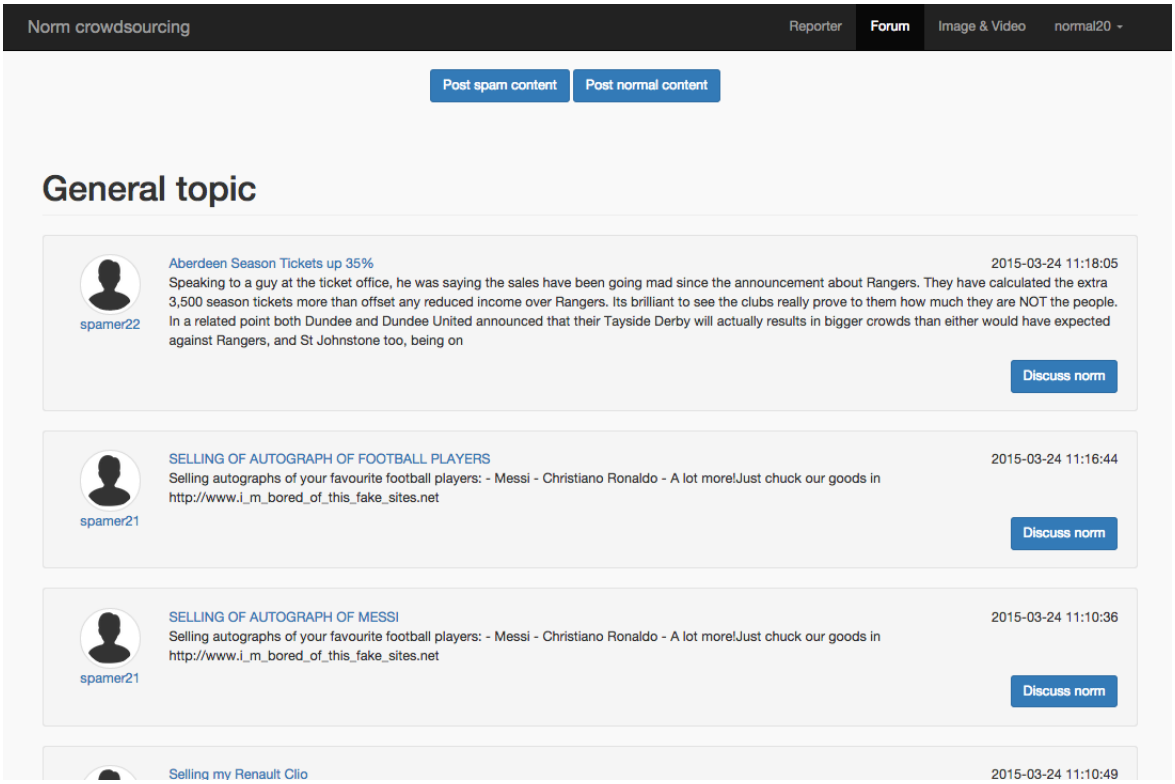


Figure 4.2: Forum section

Yii2

Yii2 [1] is a PHP framework for developing Web 2.0 applications. As a framework, it has a several interesting features like MVC, DAO/ActiveRecord, I18N/L10N, caching, authentication and role-based access control, scaffolding, testing, etc. It can reduce your development time significantly. Initially, the learning- curve is high but when you are used to use the framework it allows to program very quickly. It also has a huge community that provides some scripts or modules with extended functionalities and a API very well described.

[1] Yii2 http://www.yiiframework.com

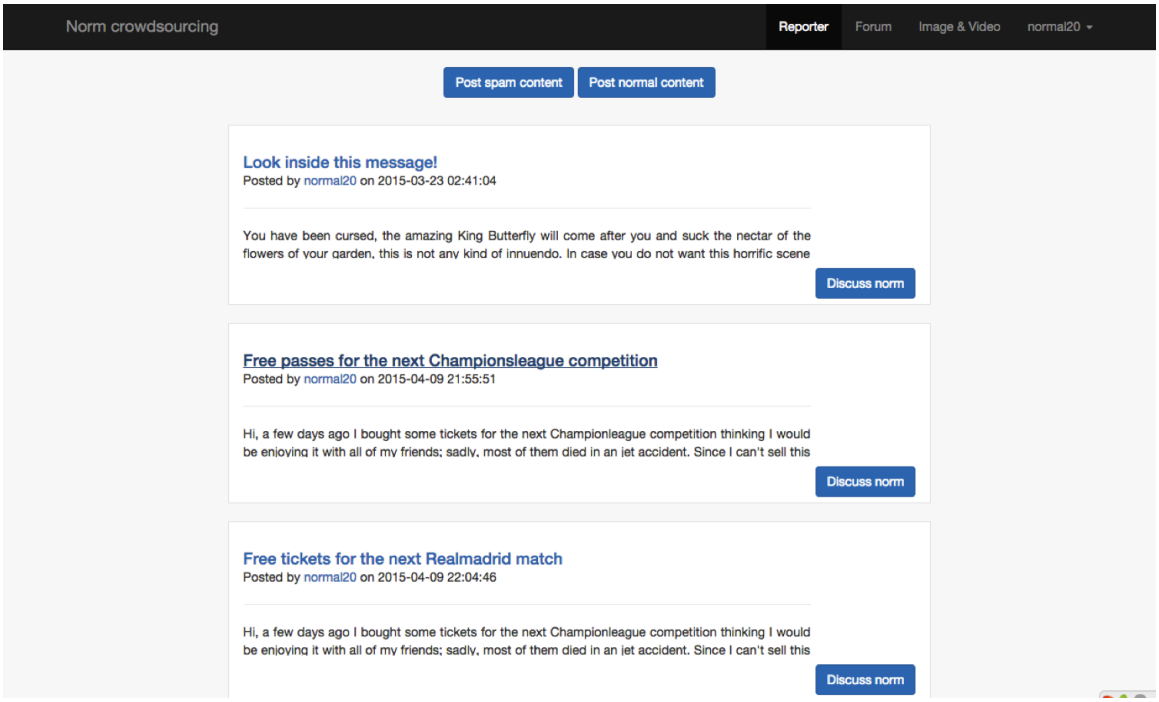


Figure 4.3: Reporter section

Xampp

XAMPP [2] is the most popular PHP development environment because it is a completely free, easy to install Apache distribution containing MySQL, PHP, and Perl. Moreover, it has version to Linux, Windows and Mac Os. Finally, it has a lot of packages with some implemented functionalities but probably one of the most interesting one is the phpmyadmin package. Phpmyadmin is one of the most popular web applications to manage the MySQL datasets.

Database

The database was developed using MySQL. It has several entities with a lot of connections between them. Figure 4.8 shows its actual structure.

Related with the dataset, one of the most interesting features that Yii2 has is related to the migrations. The migration is a set of actions in the dataset that is need to achieve some specific version of the website. The

[2]Xampp https://www.apachefriends.org/index.html

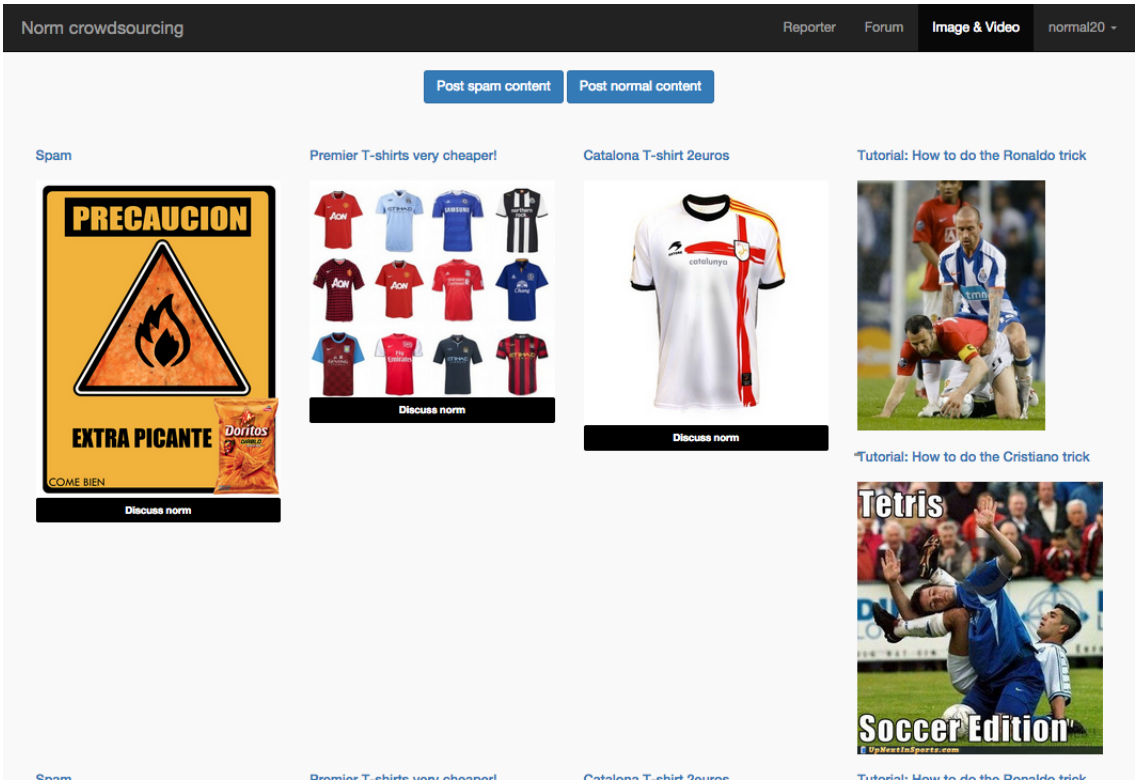


Figure 4.4: Image/Videao section

implemented website has a migration file that when it is called can create and add content to the dataset automatically.

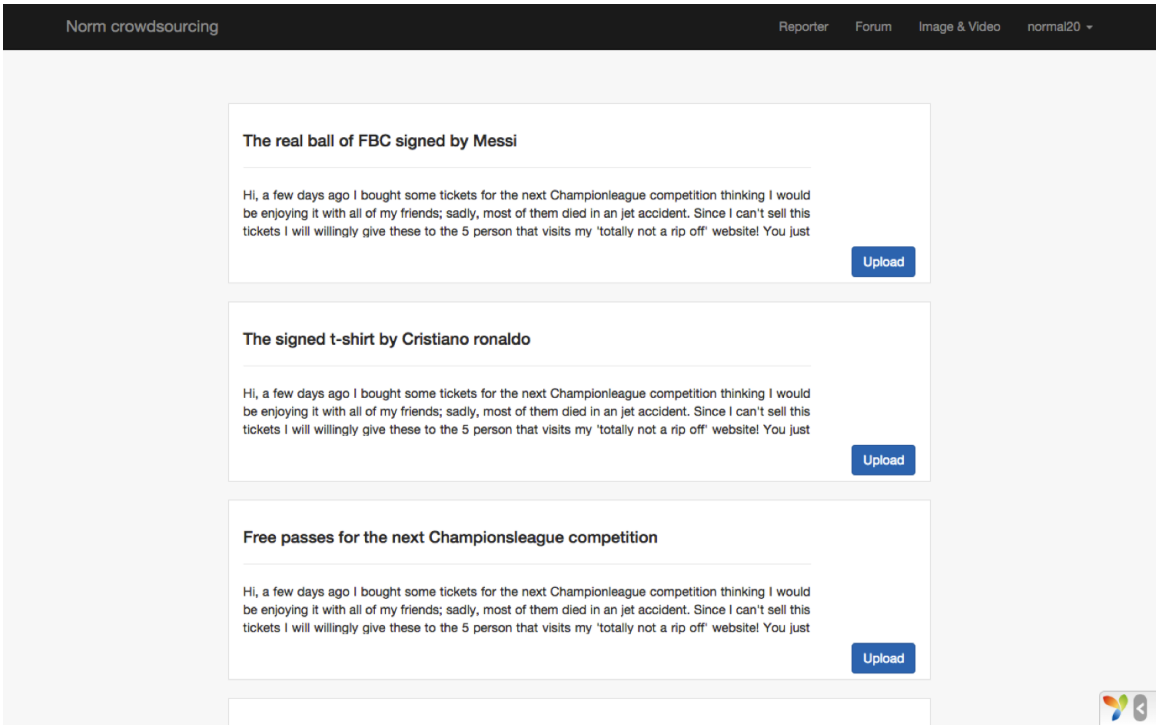


Figure 4.5: Upload content

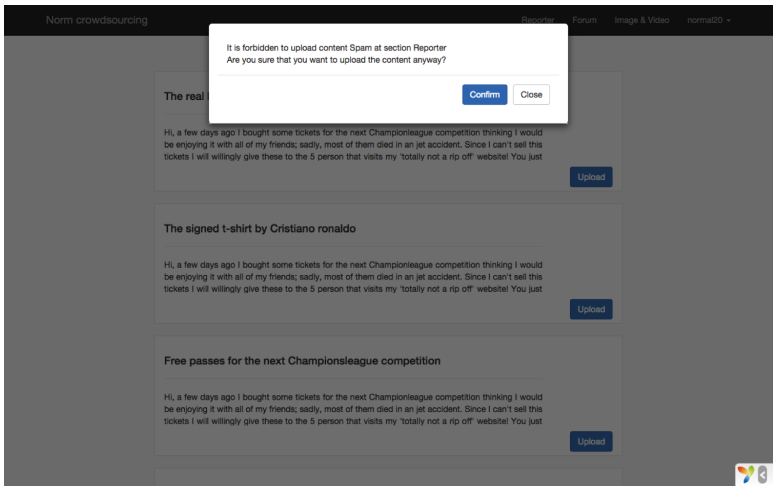


Figure 4.6: Alert triggered the user tries to upload prohibited

Figure 4.7: Active norm

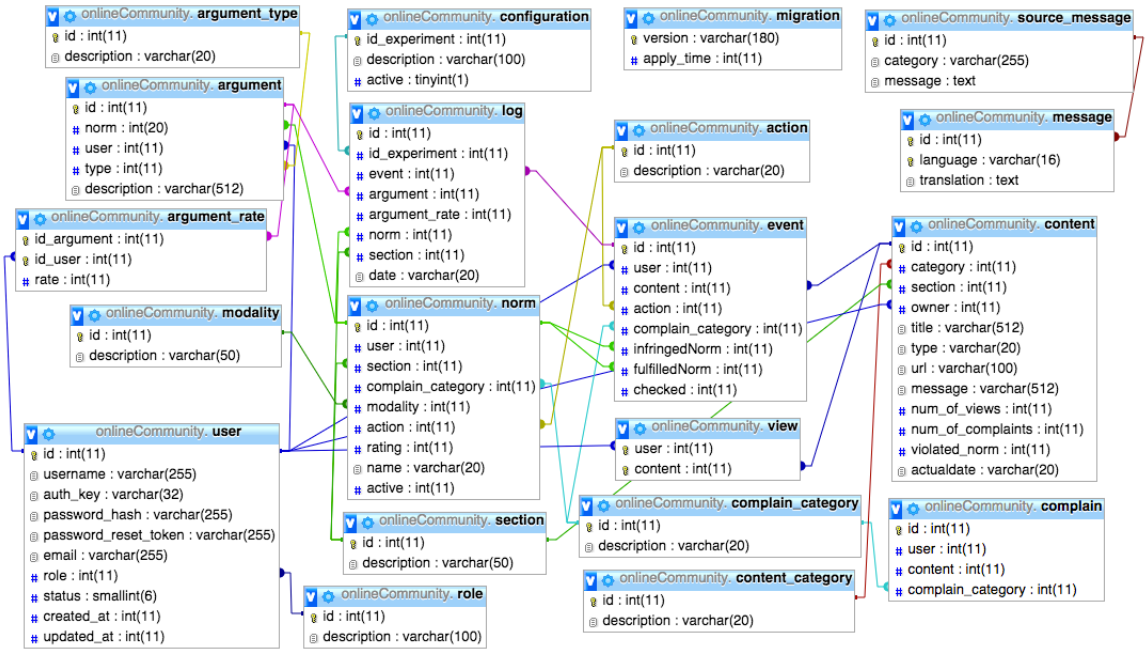


Figure 4.8: Dababase table structure

Chapter 5

Evaluation

The aim of this evaluation chapter is to study if the proposed method of norm co-creation in virtual communities works property. Furthermore, so as to test if every proposed goal was achieved, we will also evaluate if the users feel comfortable with it. Moreover, we want to study how norms are generated depending on the user types that the community has.

Preeliminary tests have been conducted with a limited number of real users that participate in two different experiments by enacting different roles. The overall objective is to assess which types of norms are generated when the virtual community is populated by a user society composed of a majority of users playing a specific role. In order to simplify, analyse and characterise better this, just two different roles were used (spammer and normal user), although in the future it could be done with more advanced experiments, such as several roles intervening in the platform or users without a pure role. Thus, although the system can deal with other types of complaints (such as pornographics, violent o rinsults), in these experiments users are restricted to just complain about spam content. It is worth noticing that this simplifying restriction limits the user interaction, since in this setting users can just regulate (or in other words, create norms and complain about) spam content, but not about any other type.

5.1 Configuration

So as to see how the platform works with people of different nationalities and genders, two experiments have been defined. Considering that the experiments have been conducted with a small number of users, every user had to enact both spammer and normal user roles in order to extract useful and variate information.

At the very beginning of the test, human testers were provided with a presentation document (appendix [6.1]). This document is not only a document which guarantees that users have been provided with sufficient information about the website and their roles within it, but it also contains a description of the purpose of the test, role types and the platform where users had to interact.

Afterwards, testers received a description of the specific tasks they were requested to perform. Within this aim, we created two alternative versions of the same task document that were used to define the user community population for each experiment (see appendix 6.1, 6.1). In this manner, in one document the human user performs the first experiment as a spam user and in the other the user participates as a normal user. Nonetheless, during the second experiment they swap roles, thus, normal users become spammers and spammers convert into normal users. These documents also contain the necessary access data, such as IP address and user credentials.

Regarding the task, we defined three rounds where users were requested to perform some specific actions with the platform, focusing subsequently in: first, uploading content; second, discussing about norms; and third, participating in the community once the norm has become active and it regulates the content.

The community population is split in two types of users (roles). On the one hand, there are users whose actions are "good" or "normal", in the sense that they do not produce conflicts when uploading their contents. Since this type of users prefer conflicts do not arise, whenever they detect conflictive contents, they promote regulations by creating and discussing about norms. Users argue about norms by adding arguments that support the norm they aim at creating. They defend norms by adding and rating high pro (positive)

Figure 5.1: Users of the test

arguments (i.e., those arguments in favour of norms).

On the other hand, spammers are users that generate conflicts by adding advertising or repetitive content. They upload content that moderate users cannot tolerate. This role or type of user tries to defend her/his content rating low the positive arguments of the norm and also by adding negative (cons) arguments to defend her/his content.

As previously mentioned, we tested the web application with 12 users (10 men and 2 women) that participated simultaneously in the test session. Specifically, Table 5.1 provides a short description of their main traits, that is, gender, age, nationality, and their actual participation in real virtual communities (in a scale of 1 to 5). From now on, we identify users with a identification number as the figure [5.1] shows. In order to anonymize the users we associate a different avatar to each user where the only realistic trait is the gender, the other characteristics of the avatar do not correspond to the real users.

Figure [5.2] illustrates the two conducted experiments. Experiment 1 is characterized by a partition of 33'3% good behaving users and 66'6% of spammers. Conversely, experiment 2 reversed this and considered a partition of 66'6% of good users and 33.3% spammers. In the first experiment the users ID1, ID2, ID3 and ID11 have the role of the normal or good user. So, their behaviour is requested to be as this role defines: creating norms when the spam users upload advertising or spam content, adding and voting positive (pro) arguments. In the opposite role the users ID4, ID5, ID6, ID7, ID8, ID9, ID10, and ID12 have the spam role. So, their defend their advertising and spam content of good users.

	Gender	Age	Nationality	Participation
	M	28	Venezuelan	3
	M	28	Spanish	4
	F	23	Hungarian	3
	F	41	Spanish	5
	M	33	Spanish	5
	M	27	Venezuelan	3
	M	22	Spanish	2
	M	24	Spanish	3
	M	23	Spanish	4
	M	24	Spanish	2
	M	23	Spanish	1
	M	24	Andorran	2
Median		24		3
Max		41		5
Min		22		1
Average		26,6		3,08

Table 5.1: General traits of human testerss

In the second experiment the roles are changed and ID1, ID2, ID3, and ID11 were requested to behave as a spammer users and ID4, ID5, ID6, ID7, ID8, ID9, ID10, and ID12 are requested to behave as a normal users.

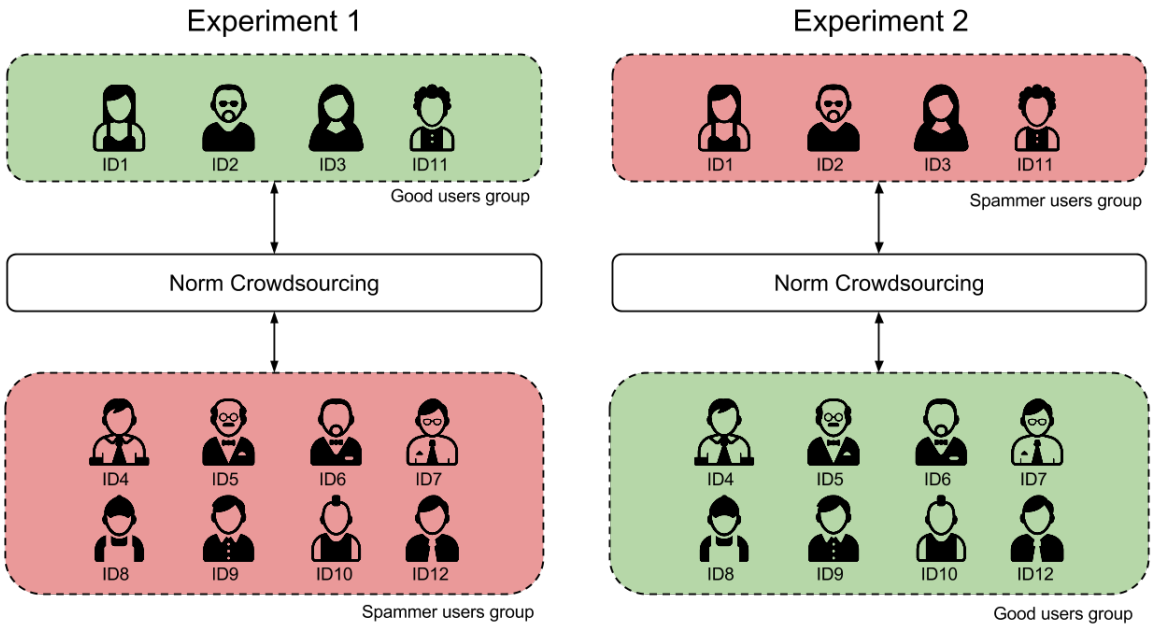


Figure 5.2: Experiments distribution

5.2 Norm activation criteria

Both experiments have the same configuration parameter. The a,b variables from the OWA operator (see section 3.4.2) where defined as a=0.3 and b=1 respectively. Considering a set of Norms, the system activates a norm $n \in N$ based on the following formulae:

$$Active(n) = \{\ true if \#DiffUsrArg >= 3 and totalRate\ = 2 false otherwise$$

The norms will be active if DiffUsrArg>=3 and totalRate>=2. #DiffUsrArg is a variable that accounts for the number of different users that added arguments to this norm. Thus, if a user adds several arguments to the norm, then the system will only compute that as one. Moreover, *totalRate*

is a variable that stands for the overall rate computed for the norm. This computation follows the specification in Subsection 3.4.2. Figure 5.5 shows an example of an active norm whose overall rate is higher than 2.

5.3 Experiments

This section details the two experiments that were conducted in this preliminary evaluation.

5.3.1 Experiment 1

As a result of the first experiment, four norms were generated. The norm 5.3, focused the discussion of the first experiment. As it can be seen, it is composed by a long discussion with several arguments and rates. Surprisingly, despite the fact that in this experiment the proportion of spammer users is far larger than the moderate users, these moderate users are really active and thus, the overall arguments appear to be somehow balanced. Nevertheless, rates clearly show that spammers defend their arguments and try to reduce the impact of positive arguments by raking them with low votes. Moreover, spammers add a significant number of negative arguments with high rates so to argue against the creation of the norm. As a consequence, the overall number of stars that the norm gets is very low (less than one). The results of the other norms 5.4,5.5,5.6 created during first experiment are not very conclusive, since argumentation is rather poor (limited discussion with few ratings, even some arguments without rates at all).

5.3.2 Experiment 2

As previously mentioned, this second experiment was populated by a higher proportion of moderate (normal) users. A result of the second experiment, three norms were generated. Differently of the first experiment, the discussion was dilated along several norms. All the norms have long discussions, with several rated arguments. This experiment does not have long argumented norm like norm 1 from the first experiment 5.3. It may be worth noticing the difference in the number of overall stars that the norms get for this experiment. Norm in Figure 5.9 does not have negative arguments and all positive arguments get a high rating, thus the resulting rating for the norm is the highest possible (that is, 5). On the contrary, norm in Figure 5.8 gets no stars in the overall ranking due to a majority of positive arguments with zero votes. Finally, Figure 5.7 shows a norm whose overall ranking is also very low, but in this case, it is due to a large number of negative rankings that are ranked quite high and that compensate the positive arguments.

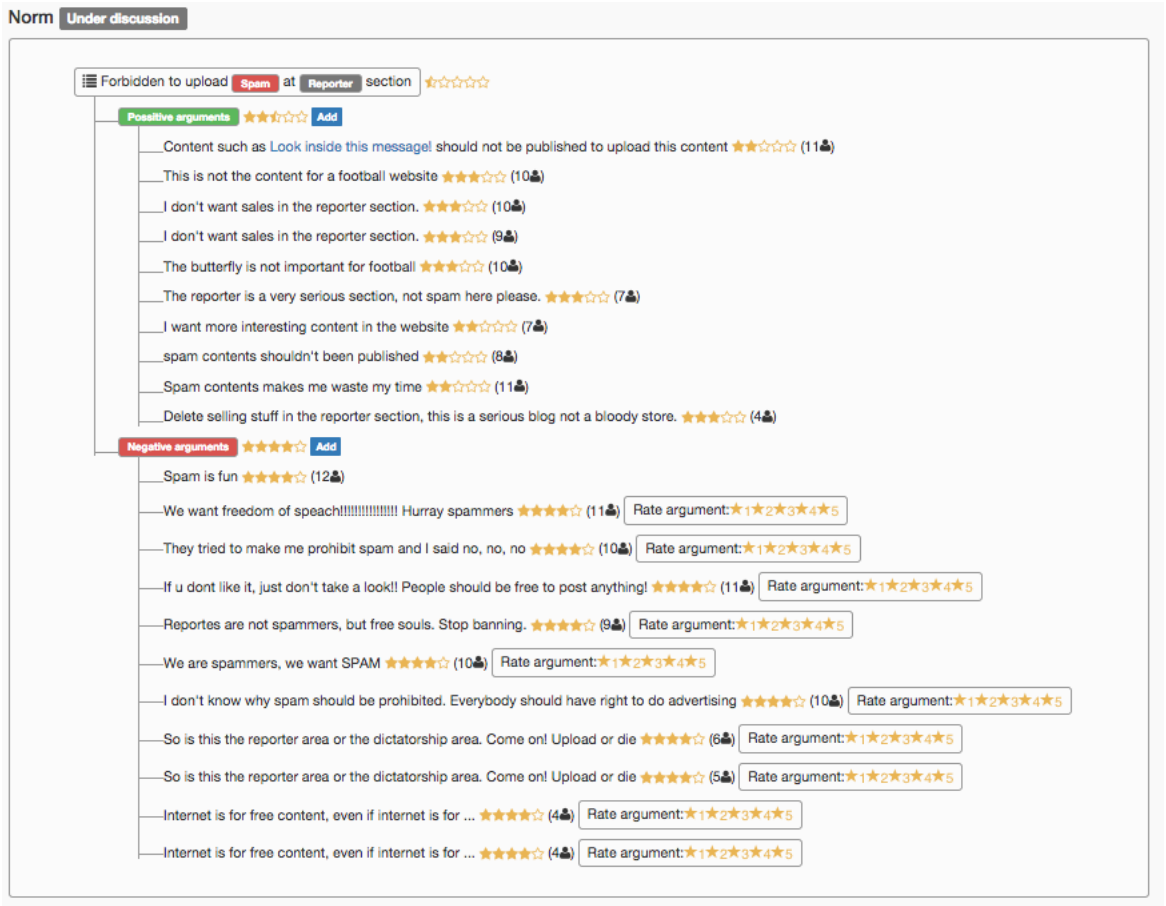


Figure 5.3: Norm 1 from first experiment

5.3.3 Results

In order to track all interactions of the users, the system is endowed with a log system that helps to monitor the actions performed by the users. It also has the possibility to download the log in several file types like excel in order to perform an efficient post-analysis. In this case, human testers did not behave exactly as expected since, for example, they did not distribute themselves evenly in the discussion of all the proposed norms. Although the configuration of the norm activation may have affected their behaviour, they still acted in different ways for different norms: some testers requested to act as moderate users did not voted for positive arguments (thus, they did not

ID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
1	3	4	5	4	5	5	4	4	4	4	2	3	5	3
2	4	1	4	4	5	5	4	3	2	2	2	5	4	2
3	3	4	4	2	5	5	4	4	5	4	3	3	4	2
4	5	1	3	1	5	5	4	4	3	3	1	4	4	4
5	5	2	4	5	5	5	5	3	5	4	2	5	5	4
6	3	2	4	4	5	5	2	4	2	4	4	5	4	3
7	2	3	3	1	4	2	2	3	1	4	4	3	4	3
8	3	2	4	5	5	5	5	5	5	5	4	5	5	3
9	4	4	5	1	3	5	5	2	4	5	1	4	5	2
10	2	2	5	2	5	4	5	3	2	1	1	4	5	2
11	1	2	4	4	3	5	3	2	2	4	4	3	4	5
12	2	1	4	4	4	5	4	3	1	3	3	4	4	4

Table 5.2: Questionnaires' results

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Median	3	2	4	4	5	5	4	3	2,5	4	2,5	4	4	3
Max	5	4	5	5	5	5	5	5	5	5	4	5	5	5
Min	1	1	3	1	3	2	2	2	1	1	1	3	4	2
Average	3,083	2,333	4,083	3,083	4,5	4,667	3,917	3,333	3	3,583	2,583	4	4,417	3,083

Table 5.3: Statistical results

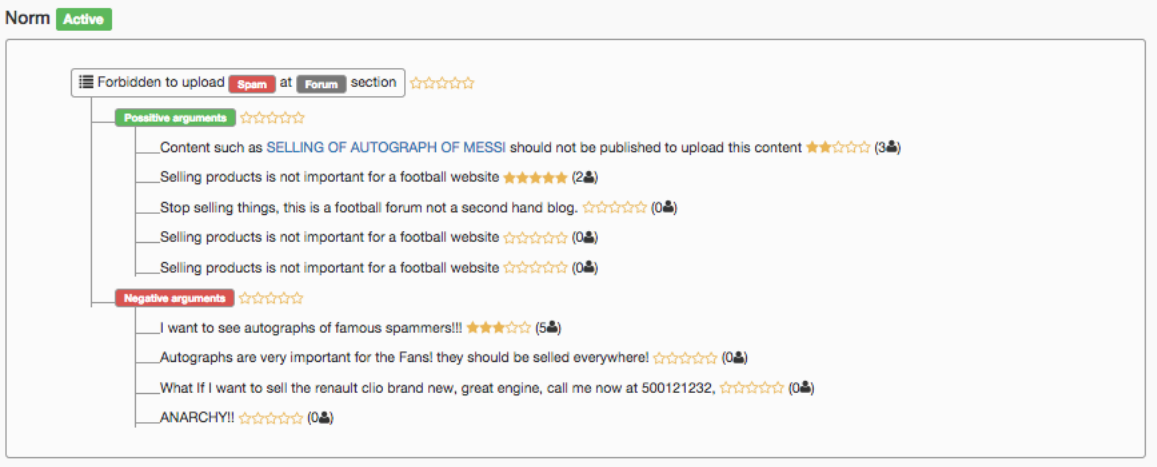


Figure 5.4: Norm 2 from first experiment

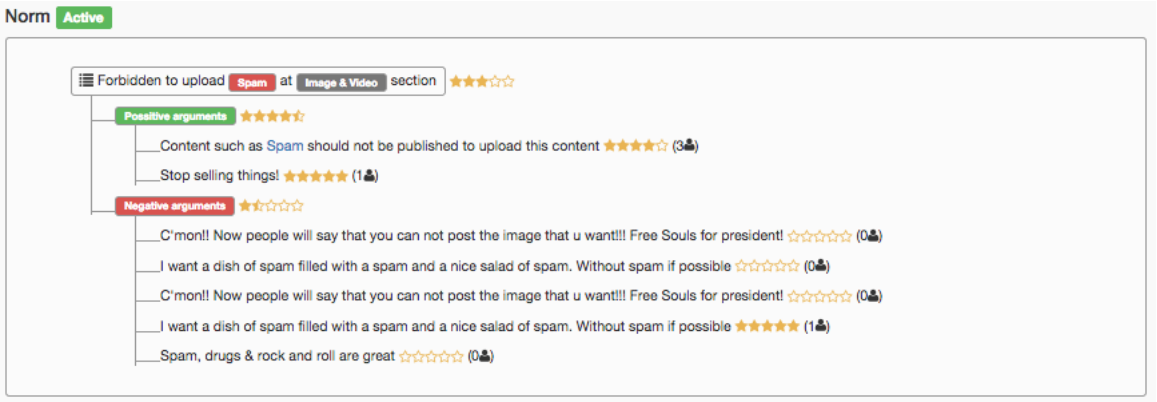


Figure 5.5: Norm 3 from first experiment

defend the inclusion of norms that prohibited spam content); some testers acting as spammers did not include negative arguments in the norms but rated negatively the norms; and so on. This Master Thesis has proposed a method for aggregating opinions about norms, and, although the actual computation provided results that make sense, it was difficult to foresee how users will actually behave when providing such opinions. Thus, this requires a further study that is out of the scope of this work.

After performing the two experiments, a satisfaction survey was provided to the human testers. Although the actual document is included in appendix

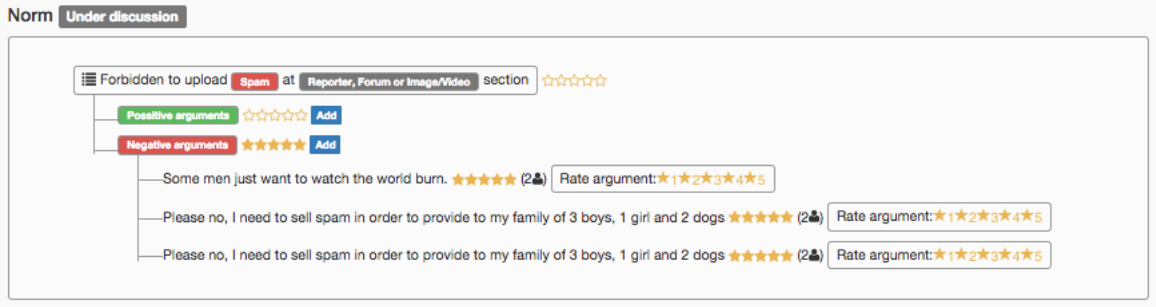


Figure 5.6: Norm 4 from first experiment

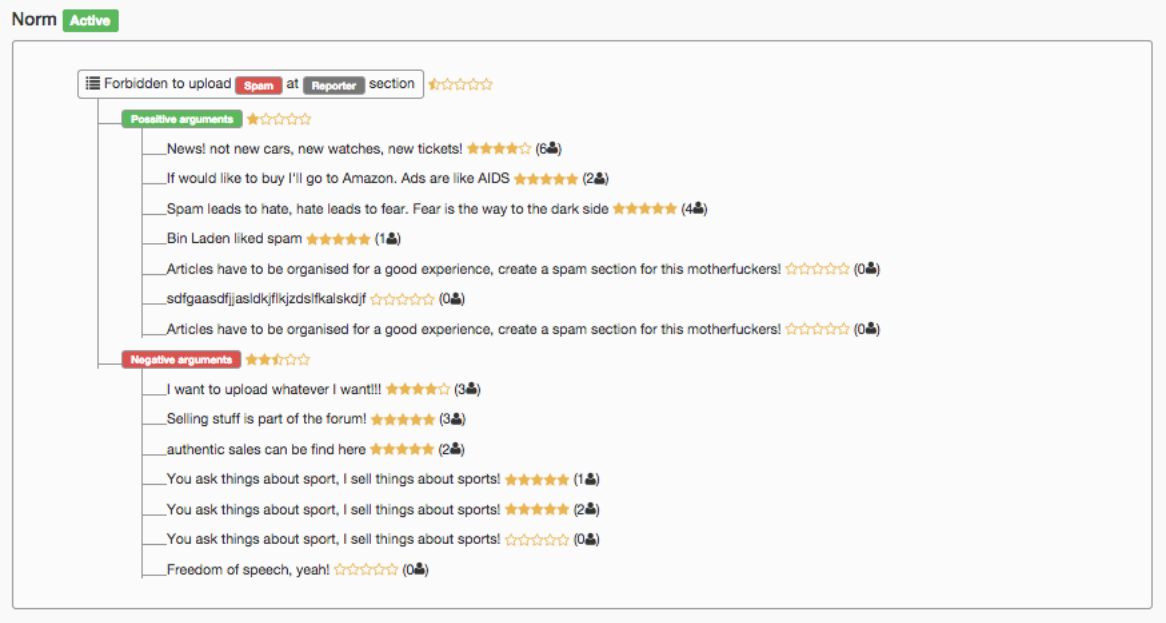


Figure 5.7: Norm 1 from second experiment

6.1, Table 5.10 summarises these questions and its results are showed in the tables [5.2, 5.11]. The survey is composed by 14 questions whose possible answers are provided as a numerical five-point Likert scale. Questions are formulated so that 1 corresponds to the most negative answer and 5 to the most positive as done in [1]. To sum up, in this survey, one means that disagreement with this question and five that total agreement with the question. For instance in the question 1: *Do you participate in virtual (social)*

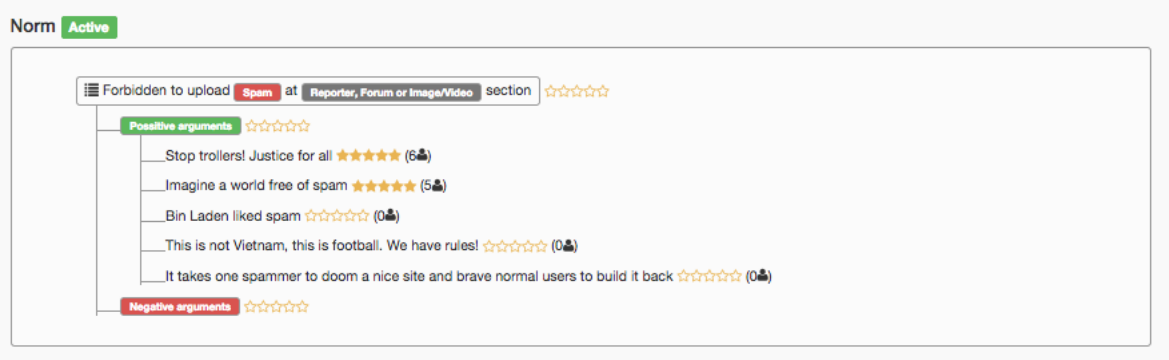


Figure 5.8: Norm 2 from second experiment

Figure 5.9: Norm 3 from second experiment

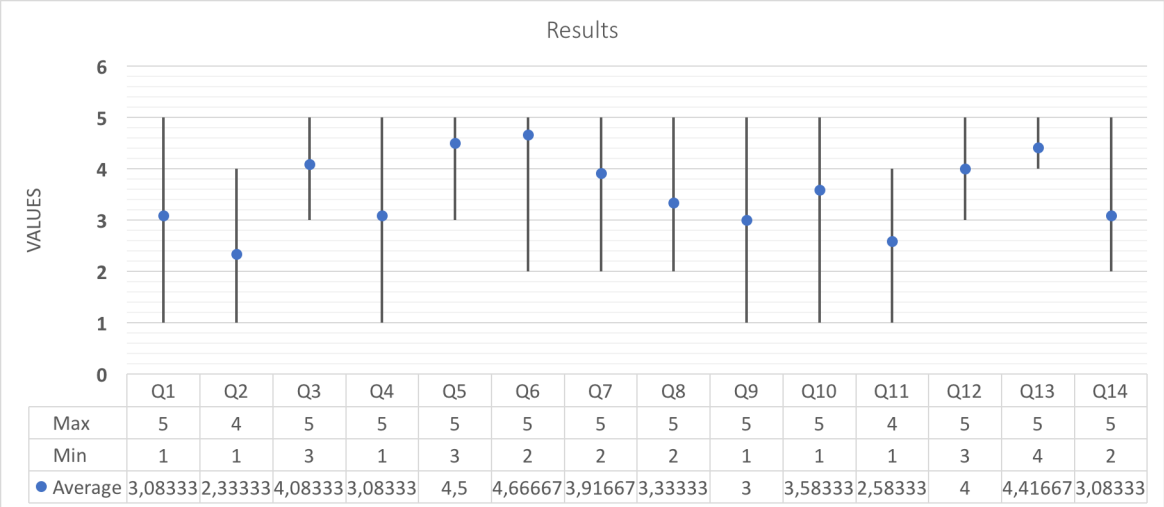
communities? one refers that your never participate and five you participate a lot in virtual social communities. In the middle, could be sporadic users that enter in the virtual social communities not very often. Having a look at the questions, one can easily notice that some questions are linked between them such as questions Q5 and Q6, Q7 to Q9, Q11 to Q13. Q5 and Q6 are meant to assess if some type of contents annoy the users. Furthermore, questions ranging from Q7 to Q9 are related with the moderation necessity and try to quantify if users would participate in this process. Finally, questions ranging from Q11 to Q13 are related to the punishments for those users that not follow the norms of the system. Thus, these three questions capture if punishments are actually required and, in case they are, if they should also

Question	Description
1	Do you participate in virtual (social) communities?
2	Are you aware of the participation policies (i.e., norms) that social communities such as Facebook or Twitter have?
3	Do you think norms are useful in social communities?
4	Have you ever felt uncomfortable when participating in a social network?
5	Do you think spam content is annoying?
6	Do you agree that some content types (e.g., fake, rude, porn, violent content) should be prevented from appearing in a social community?
7	Do you think moderation is necessary in a social community
8	Do you prefer to empower users to create their own norms rather than having moderators?
9	Would you be active in the creation of norms in a social community?
10	Did you easily follow the discussions on norms during the experiments?
11	When playing spammer role and uploading spam, did the norm infringement alarm make you feel uncomfortable?
12	Did you miss the application of punishments to the people that did not follow the norms?
13	Do you think it is necessary to discuss about both punishments and norms?
14	How usable was the website?

Figure 5.10: Questions

be discussed by means of the same argumentation process as norms.

Table 5.2 summarises the opinion of the human testers as reflected in their



	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Max	5	4	5	5	5	5	5	5	5	5	4	5	5	5
Min	1	1	3	1	3	2	2	2	1	1	1	3	4	2
Average	3,08333	2,33333	4,08333	3,08333	4,5	4,66667	3,91667	3,33333	3	3,58333	2,58333	4	4,41667	3,08333

Figure 5.11: Box plot

answers to the satisfactory test. In addition, Table 5.3 and Figure 5.11 show the statistical values associated to the different opinions for each question.

Along the following items we analyse the results obtained for each question:

- *Q1: Do you participate in virtual (social) communities?* Answers to this question show a wide variation in the profile of testers regarding their involvement to regular on-line communities. Thus, some testers (ID10, ID11 and ID12) don't participate frequently in the virtual (social) communities whereas the rest of testers do actually do it.

- *Q2: Are you aware of the participation policies (i.e norms) that social communities such as Facebook or Twitter have?* All answers to this question show that the users don't know the norms. This problem may arise from actual websites because the norms there are not properly showed nor explained. Usually, this part of the websites is presented in the form of text-free regulations described as long texts that people hardly devote the time to read. Thus, norms clarity might benefit from providing structure to norms.

- *Q3: Do you think norms are useful in social communities?* Answers do have a high rate. This means that users perceive that the norms are

needed in order to avoid conflicts. Also, this question gains importance if you link its results with those from Q2. The testers are not aware of the norms of their social communities but still think that are useful.

- *Q4: Have you ever felt uncomfortable when participating in a social network?* This question also presents a wide range of answers, since the minimum value is 1 and the maximum is 5. This may due to the fact that the uncomfortable feeling depends a lot on previous experience. If any tester suffered a bad experience he/she will answer in this question as if he/she felt uncomfortable at some point.

- Both questions *Q5: Do you think spam content is annoying?* and *Q6: Do you agree that some content types (e.g, fake, rude, porn, violent content) should be prevented from appearing in a social community?* have a high rate in their answers. Overall, users feel that some type of content is annoying. This is also related to the answers to Q5 where users perceive norms as being useful. Therefore, in order to avoid this annoying content it is necessary to add norms to the system. It is also worth noticing that most testers consider spam content to be less annoying than other types. The test was concluded with spam content to avoid confronting testers with inappropiate content that may offend them, but question Q6 was included to assess that indeed other types of contents are most in need of regulation.

- *Q7: Do you think moderation is necessary in social communities?*, *Q8: Do you prefer to empower users to create their own norms rather than having moderators?* and *Q9: Would you be active in the creation of norms in a social community?* reflect that some users (ID3, ID5 AND ID8 most specially) will participate in the creation of the norms. However, ID12, ID7 will not participate in this process although they admit moderation is necessary.

- *Q10: Did you easily follow the discussions on norms during the experiments?* has an average answer above 3,58 which can be interpreted as users could follow the discussion on norms in seamless way. Just two users (ID2 and ID10) reported a low answer, and this may be caused by some technical difficulties suffered during the test. Moreover, this result validates the aim 1,2 of the project that proposes to define a

method to support users with the co-creation of their own norms and to built a system in which actual users can co-create their own norms.

- *Questions Q11: When playing spammer role and uploading spam, did the norm infringement alarm make you feel uncomfortable?*, *Q12: Did you miss the application of punishments to the people that did not follow the norms?* and *Q13: Do you think it is necessary to discuss about both punishments and norms?* are asking about the punishments. Punishments are actions when an user does not follow the rules of the system. Some common punishments in on-line social communities are removing the content or banning the user that uploaded it. The proposed solution shows an alert whenever an user tries to upload content but exists a norm that regulates this type of content. From the answers to question 11 we can conclude that the alarm is not enough to avoid these types of behaviour. Nevertheless, results from Q12 and Q13 show that users consider that it is both necessary to include punishments to norms as well as to discuss about them.

- Regarding the usability of the system, the question *Q14: How usable was the website?* has got a mean value around 3. This value is above the middle value, and thus can be considered as moderately positive. Nevertheless, in favour of the system, it can be taken into account previously mentioned technical problems (our local server did not support properly the simultaneous connections of the testers and sometimes they experienced lag) did affect user's experience, and this somehow biased the results obtained for this question. So, this result would be consider as a non-indicative although it is more that the middle value.

Chapter 6

Conclusions and Future work

A new way to manage, control and discuss norms has been proposed and implemented in this Master Thesis. The *Norm argument map* implemented guarantees that users can follow up the discussion about the norms and can easily rate, and aggregate new arguments. The norm structure in the form of a tree helps users to give and share their opinions. Although the interaction with the structure is focused on norm arguments instead of focusing on the norms themselves, users still have the impression they discuss about norms. To guarantee that users do not lose the feeling of being discussing about norms is key for the norm co-creation process, even if users add their opinions in the form of arguments and argument ratings.

The new way to manage, control and discuss the norms has been developed in a web application where the users can manage, control and discuss their own norms. In addition, with its log facility, it is possible to export in several types that is very useful to monitor the users within the platform.

The results of the experiments show that people tend to focus on the discussion about specific norms. When a norm has a long discussion people tend to participate in the discussion. However, if the norm does not have a long discussion it is more unlikely that the users join it.

User preferences and opinions in the satisfaction survey vary widely from person to person, since each person may experience in a unique way his/her interaction in the social community and this greatly determines his/her resulting opinion. Moreover, the users report not to be aware of current norm

systems in virtual social communities like Twitter or Facebook. However, they acknowledge that norms are necessary to prevent some undesirable behaviours. Furthermore, users demand punishments as a component that should be added to the proposed method. In addition, users like the system and the way that they can propose and discuss about norms. The results of the tests show that the aims of the project were achieved at least up to a certain level, since the majority of users reported they liked the norm discussion proposed in the application.

6.1 Future work

As future work, some points that deserve improvement are be the following:

- Information aggregation: The OWA operator was a success when aggregating different arguments of a norm. The implementation is easy and the computational cost is low. Nevertheless, it could be interesting as a future work to compare different information aggregation algorithms to see the impact of this aggregator operators in the final output of the norm.

- Population size: The main problem when dealing with testing with humans is to recruit them. In this case the problem gets worse because all users must be simultaneously participating in the social community and, for legal reasons, they were also required to be physically in the testing place. One of the greatest problems to deal with while testing is related to the number of recruited testers. Is not an easy task to recruit a huge volume of testers for the application because the date and place could be an impediment. So, in order to solve this problem, one possibility could be do the tests in different dates and places, covering different towns or zones. Also, one solution to increase the number of human testers could be use services as Amazon Mechanical Turk (AMT) [1] where they provide a infrastructure where you publish a remunerate task and the users could test your task.

- Content generated by the users with a content classification component where it guarantees the correct labelling of the content could be really

[1] Amazon Mechanical Turk, https://www.mturk.com/mturk/welcome

useful extra component. Now, the platform can only deal with predefined content that was generated and labelled beforehand. However, is not a typical case of a virtual social community because in this type of social webs, users generate new content. In order for the social web to admit users to upload new content, it would be required to include an advanced natural language processing component that automatically labels (classifies) content by its type (e.g., spam, violent, or adult).

- Comparing the norm argument map structure with other types of norm structures would also be an interesting line of future work. Although the results of the test show that the majority of people like this type of structure to discuss the norms because it permits to follow easily the discussion, it could be nice in the future to provide several alternative norm structures in order to see the difference in response time, feeling about the discussion, size of the discussion and so on. Also, norms could have relationships between them , for example, in the form of parent dependencies, and thus it becomes important to deal with them when considering norm arguments. In this case, many issues related to the topology of the relationships arise, since for example, adding arguments to a child norm this implies that the arguments will also belong to the parent, or alternatively, it may be more suitable to just keep it at the level of the child norm?

- Several test users request punishments to the users that do not follow the norms. The alert provided by the website is not enough. For this reason it could be nice to add more punishments like push out users, cold users, restrict their actions and so on.

Bibliography

[1] P. Almajano, M. Lopez-Sanchez, I. Rodriguez, and E. Mayas. Conversational structured hybrid 3d virtual environments. pages 63:1–63:8, 2014.

[2] E. Cables and M.T. Lamata. On weighted media aggregation. *International Journal of Uncertainty, Fuzziness and Knowledge-based systems*, 1:101–113, 1994.

[3] E Cables and M.Teresa Lamata. Owa weights determination by means of linear functions. *Mathware & Software Computing*, 16:107–122, 2009.

[4] R Fullér. Owa operators in decision making. *Exploring the limits of support systems, TUCS General Publications*, 3:85–104, 1996.

[5] J Keeshin. Social norms on the web: How to create productive digital communities, 2010.

[6] L. Xiao; W. Zhang; A. Przybylska; A. De Liddo; G. Convertino; T. Davies; M. Klein. Design for online deliberative processes and technologies: Towards a multidisciplinary research agenda. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 865–868, 2015.

[7] M. Klein. Enabling large-scale deliberation using attention-mediation metrics. *Computer Supported Cooperative Work*, 21(4-5):449–473, 2012.

[8] A Litvinenko. Social media and perspectives of liquid democracy: The example of political communication in the pirate party in germany. *12th European Conference on eGovernment (ECEG 2012)*, 3:403–407, 2012.

[9] J. Morales, I. Mendizabal, D. Sanchez-Pinsach, M. Lopez-Sanchez, and J.A. Rodriguez-Aguilar. Applying iron to a virtual community scenario. pages 316–325. IOS Press, 2013.

[10] J. Morales, I. Mendizabal, D. Sanchez-Pinsach, M. Lopez-Sanchez, and J.A. Rodriguez-Aguilar. Using iron to build frictionless online-communities. *Artificial Intelligent Communications*, (1):55–71, January 2015.

[11] J. Morales, D. Sanchez-Pinsach, M. Lopez-Sanchez, I. Mendizabal, J.A. Rodriguez-Aguilar, M. Wooldridge, and W. Vasconcelos. Extending normlab to spur research on norm synthesis. In *Proceedings of the fourteenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS2015)*, pages 1931–1932. IFAAMAS, 2015.

[12] D. Sanchez-Pinsach. Moderación automática en comunidades virtuales. Master's thesis, Final degree project of BsC on Computer Science on University of Barcelona, 2013.

[13] V. Torra and Y. Narukawa. Modeling decisions: Information fusion and aggregation operators. 2007.

[14] M Wooldridge. *An Introduction to MultiAgent Systems*. 2009.

[15] Ronald R Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *Systems, Man and Cybernetics, IEEE Transactions on*, 18(1):183–190, 1988.

[16] Ronald R Yager. An overview of methods for determining owa weights. *International Journal of Intelligent Systems*, 20(8):843–865, 2005.

Acronyms

DAO Data Access Object. 23

I18N Internalization. 23

L10N Localization. 23

MVC Model View Controller. 23

OWA Ordered Weighted Aggregator. 9, 19

QA Question-answers systems. 12, 18

Appendices

PRESENTATION

Norm crowdsourcing

Introduction

You are participating in an experiment to evaluate *norm crowdsourcing*, a prototype of a social community website that we have developed. This document contains a brief description of the website. We would like to stress that we are evaluating *norm crowdsourcing* and not the tester (this is not an exam you should pass). We encourage you to **be aware of your thoughts and impressions** along the test, since you will be asked about your opinion by completing a survey at the end of this test.

Norm crowdsourcing is a website focused on football content where the user (a football fan) can: upload predefined contents; view contents and complain about contents uploaded by others; and create and discuss about the virtual community norms. The website is structured in three main sections ("The reporter", "Forum" and "Image & video") that show contents in different ways but support the same user interaction (users can perform the same actions for all three sections).

The goal of this web is to allow users to decide and manage participation rules required to keep a good atmosphere in the website, so all the participants feel comfortable with the published contents and no additional moderation is required. For the sake of test simplicity, users cannot create new content: they can just upload contents from a predefined database.

Within *Norm crowdsourcing* website, users discuss about participation rules (norms) by providing and rating arguments in favour and against them. Once a norm has enough support by the user community, it becomes active and participants are asked to comply with them.

Roles

We consider two types of users within this test of the website:
Good user: typical user that uploads content that does not lead to conflicts (i.e., complaints) and promotes the creation of rules that avoid the subsequent upload of conflictive content on the site. The actions performed by good users are:

- Upload appropriate (i.e., non-spam) content.
- View content uploaded by others.
- Complain about inappropriate (i.e., spam) content.
- Promoting norms that aim at preventing spam content by adding and rating positively arguments that support them.
- Comply with active norms.

Spammer user: user that generates a lot of advertising (spam) content and tries to protect this type of content by never complaining about it (even if uploaded by other users) and by opposing to the creation of rules that prohibit spam content. The actions performed by spammers are:

- Upload both spam and appropriate content.
- View content uploaded by others.
- Opposing to norms that aim at preventing spam content by adding and rating positively arguments that go against them.
- Infringe active norms that regulate spam.

TASK good- spammer

Norm crowdsourcing

Introduction

You are participating in a test for evaluating *norm crowdsourcing*, a prototype of a virtual (social) community where users define their own participation rules (norms).

During this test, you will be asked to participate in two subsequent tests composed by the same scheduled activities:

- First test: You will be expected to act as a "good" user, namely as a user that uploads contents that is not found bothering by other users.
  - **Login** to http://161.116.52.177/Website/socialNorms/frontend/web/index as **normal1X**
  - 10 minute-long period participating in the social community:
    - Visit the "Reporter", "Forum" and "Image & Video" sections and view (click on) content there.
    - Complain about the spam contents you view.
    - Upload appropriate content.
    - You may discuss about some norms associated to contents
  - 10 minute-long period for discussing about norms:
    - Go to the "Norms" area (user menu> Norms option)
    - There, add and rate arguments **in favour** of norms that prohibit spam to be uploaded.
  - 5 minute-long period participating in the social community (now regulated):
    - Visit the "Reporter", "Forum" and "Image & Video" sections and view (click on) content there.
    - Upload appropriate content.
  - Logout

- Second test: You will be expected to ask as a spammer.
  - **Login** to http://161.116.52.177/Website/socialNorms/frontend/web/index as **spamer2X**
  - 10 minute-long period participating in the social community:
    - Visit the "Reporter", "Forum" and "Image & Video" sections and view (click on) content there.
    - Upload both spam and appropriate content.
    - You may discuss about some norms associated to contents
  - 10 minute-long period for discussing about norms:
    - Go to the "Norms" area (user menu> Norms option)
    - There, add and rate arguments **against** norms that prohibit spam to be uploaded.
  - 5 minute-long period participating in the social community (now regulated):
    - Visit the "Reporter", "Forum" and "Image & Video" sections and view (click on) content there.
    - Upload both spam and appropriate content.
  - Logout

TASK spammer- good

Norm crowdsourcing

Introduction

You are participating in an experiment to evaluate *norm crowdsourcing*, a prototype of a virtual (social) community where users define their own participation rules (norms).

During this experiment, you will be asked to participate in two subsequent tests composed by the same scheduled activities:

- First test: You will be expected to ask as a spammer.
  - **Login** to http://161.116.52.177/Website/socialNorms/frontend/web/index as **spamer1X**
  - 10 minute-long period participating in the social community:
    - Visit the "Reporter", "Forum" and "Image & Video" sections and view (click on) content there.
    - Upload both spam and appropriate content.
    - You may discuss about some norms associated to contents
  - 10 minute-long period for discussing about norms:
    - Go to the "Norms" area (user menu> Norms option)
    - There, add and rate arguments **against** norms that prohibit spam to be uploaded.
  - 5 minute-long period participating in the social community (now regulated):
    - Visit the "Reporter", "Forum" and "Image & Video" sections and view (click on) content there.
    - Upload both spam and appropriate content.
  - Logout

- Second test: You will be expected to act as a "good" user, namely as a user that uploads contents that is not found bothering by other users.
  - **Login** to http://161.116.52.177/Website/socialNorms/frontend/web/index as **normal2X**
  - 10 minute-long period participating in the social community:
    - Visit the "Reporter", "Forum" and "Image & Video" sections and view (click on) content there.
    - Complain about the spam contents you view.
    - Upload appropriate content.
    - You may discuss about some norms associated to contents
  - 10 minute-long period for discussing about norms:
    - Go to the "Norms" area (user menu> Norms option)
    - There, add and rate arguments **in favour** of norms that prohibit spam to be uploaded.
  - 5 minute-long period participating in the social community (now regulated):
    - Visit the "Reporter", "Forum" and "Image & Video" sections and view (click on) content there.
    - Upload appropriate content.
  - Logout

SATISFACTION SURVEY

Norm crowdsourcing

Participants' name _____ Signature_____

Please answer the questions below considering 1 as very low, 2 as low, 3 as medium, 4 as high, and as 5 very high

1. Do you participate in virtual (social) communities?				
1	2	3	4	5

2. Are you aware of the participation policies (i.e., norms) that social communities such as Facebook or Twitter have?				
1	2	3	4	5

3. Do you think norms are useful in social communities?				
1	2	3	4	5

4. Have you ever felt uncomfortable when participating in a social network?				
1	2	3	4	5

5. Do you think spam content is annoying?				
1	2	3	4	5

6. Do you agree that some content types (e.g., fake, rude, porn, violent content) should be prevented from appearing in a social community?				
1	2	3	4	5

7. Do you think moderation is necessary in a social community?				
1	2	3	4	5

8. Do you prefer to empower users to create their own norms rather than having moderators?				
1	2	3	4	5

9. Would you be active in the creation of norms in a social community?				
1	2	3	4	5

10. Did you easily follow the discussions on norms during the experiments?				
1	2	3	4	5

11. When playing spammer role and uploading spam, did the norm

infringement alarm make you feel uncomfortable?				
1	2	3	4	5

12. Did you miss the application of punishments to the people that did not follow the norms?				
1	2	3	4	5

13. Do you think it is necessary to discuss about both punishments and norms?				
1	2	3	4	5

14. How usable was the website?				
1	2	3	4	5

What did you like?, what do you think is missing?, ideas or suggestions are welcome…