Memòria del TFC

Umbra Filtre Web:
Aplicació de filtrat de pàgines web

Desenvolupament pendent d'escollir una llicència per la publicació.

Alfonso Pantoja Franco
ETIS

Consultora: Maribel March Hermo

21 de juny del 2006

Dedicatòria

Aquest projecte està dedicat als meus pares, a la meva germana i a la meva xicota per la paciència que sempre han tingut amb mi (i espero que la continuin tenint).

Resum

Segons Miniwatts Marketing Group (www.InternetWorldStats.com) entre l'any 2000 i 2005 l'ús d'Internet al món va augmentar el 183,4% mentre que a Espanya els usuaris van crèixer en un 218,2%.
Paral·lelament, el nombre de servidors d'Internet (*hosts*) va experimentar un increment, durant aquest mateix periode, de més de 109 mil·lions. Al gener de 2006 hi havia més de 394 mil·lions de servidors (font Internet Systems Consortium, www.isc.org).
La popularització d'Internet ha esdevingut una autèntica revolució i la "xarxa de xarxes" té cada vegada usuaris més joves.

La conseqüència òbvia d'aquests fets és el creixement del nombre de tot tipus de continguts i que poden ser accedits per gairebé tothom sense cap tipus de restricció motiu pel qual durant aquests últims anys ha sorgit la necessitat de protegir als menors en front de continguts no adequats per ells, i és en aquest punt on tenen cabuda els programes de filtratge de pàgines web (que es el medi més popular per accedir a la informació).

Aquests programes tenen com a finalitat restringir l'accés a diversos tipus de continguts segons uns criteris establerts i no estan orientats exclusivament als pares sino que també es fan servir a empreses i organitzacions.

El projecte presentat explora les possibilitats de la tecnologia de Microsoft .NET Framework en el camp del filtratge de pàgines web.
Per disseny s'ha escollit una solució relacionada amb xarxes d'ordinadors i el resultat del projecte és una aplicació que utiliza com a base un proxy http, de forma que és capaç de filtrar pàgines web abans de que aquestes arribin al navegador.

Index

Index de continguts

Index de taules

Index de figures

1 Introducció

1.1 Justificació del TFC i context

Actualment existeixen vàries aplicacions de filtratge web, de fet, si es realitza un procés de cerca a Internet força intensiu es poden arribar a trobar unes 30. D'aquestes aplicacions es va escollir una mostra de 9 per tal d'analitzar les seves característiques principals d'on es va concloure que la majoria d'elles només funcionen si es fa servir el navegador Internet Explorer[3], i com a conseqüència d'això no és possible cap tipus de filtratge si s'utilitzen navegadors com ara Mozilla FireFox[2] o Opera[3].

Les funcionalitats més típiques que es van trobar a tots ells són: bloqueig web per paraules clau (contingut de la web), per URL (llista negra i llista blanca), enregistrament d'activitat (webs visitades i bloquejades) i protecció del programa mitjaçant password.

Pel que s'ha observat només un d'ells sembla funcionar com un proxy HTTP (que és la opció que garantiria la total compatibilitat amb múltiples navegadors) i la resta funcionen interceptant o bé els ports (com el cas d'Internet Controller Quattro), fent crides a les llibreries de xarxa de Windows o a llibreries d'Internet Explorer i és precissament aquesta la principal motivació del projecte de l'aplicació de filtratge Umbra.

Tabla 1-1 Comparativa d'aplicacions de filtratge existents

Aplicació (autor)	Filtratge web	Bloquejos	Registre i Monitorització	Protecció privacitat	Limitació temps connexió	Altres	Requisits tècnics mínims / Observacions
NetNanny 5 (Net Nanny)[4]	- Llista negra - Llista blanca - Paraules prohibides	- Chats i Missatgeria Instantània. - Grups de notícies i jocs violents. - Descàrrega de material amb copyright i obscè. - Finestres emergents.	- Webs visitades. - Converses de chat i amb qui es parla.	SI (Filtratge de la informació que s'envia a webs, chats i emails)	- Màxim hores diàries totals. - Màxim hores per cada usuari. - Bloqueig accés a Internet per horari.	- Protecció per contrasenya.	- Windows 98, Me, NT 4, 2000, XP. - Pentium o AMD Athlon (K6 o superior). - 32MB RAM i 50 MB disc dur. - Internet Explorer.
Control Kids (Proxymis)[5]	- Llista negra (per categories temàtiques) - Llista blanca.	- Finestres emergents (són tancades automàticament).	-Webs visitades. - Intents d'accés a webs prohibides	SI	NO	- Protecció per contrasenya. - Intercepta pulsació de tecles.	-Windows 98, Me, 2000, XP. - Internet Explorer.

Aplicació (autor)	Filtratge web	Bloquejos	Registre i Monitorització	Protecció privacitat	Limitació temps connexió	Altres	Requisits tècnics mínims / Observacions
		- Banners publicitaris. - Eliminació d'Spyware. - Descàrrega de fitxers (permesa o bloquejada segons autorització).	.			- Eliminació rastres navegació.	
FreeShield 5 (Free Shield)[6]	- Llista negra - Llista blanca - Segons algorismes del programa. - Impideix buscar als cercadors usant paraules no adecuades per nens.	- Protecció contra dialers. - Detecció i eliminació d'Spyware/Ad-ware. - Finestres emergents.	- Webs visitades			- Protecció per contrasenya. - Eliminació rastres navegació, històric fitxers oberts, paperera de reciclatge.	- Windows 95, 98, Me, NT, 2000, XP, 2003. - Internet Explorer (el programa és una toolbar per aquest navegador).
Internet Controller Quattro 4.01 (Internet Controller)[7]	- Segons algorismes del programa. - Llista blanca.	- Descàrregues. - Webs extrangeres. - Finestres emergents. - P2P i Missatgeria Instantània. - Email.	- Webs visitades.	NO	NO	- Protecció per contrasenya.	- Windows. - El programa obre molts ports esperant les respostes HTTP dels servidors remots això afecta al rendiment pero funciona amb qualsevol navegador. Té algunes errades.
FreeProxy (Hand-Crafted Software)[8]	- Llista negra - Llista blanca - Segons algorismes del programa.		- Webs visitades. - Intents d'accés a webs prohibides.	NO	- Bloqueig accés a Internet per horari. - Accés a certes webs per horari.	- Protecció per contrasenya.	- Windows 98, NT, 2000, XP, 2003 Server. - Linux - Internet Explorer, Netscape, Mozilla… - Incorpora servidor web i cache HTTP per servir pàgines a una intranet de forma segura.
ChildWebGuardian 2.6 (Zecos Software)[9]	- Llista negra - Llista blanca - Paraules prohibides	- Finestres emergents (les que no tenen barra d'eines i aquelles que tenen certa paraula al títol)	- Webs visitades - Webs bloquejades i motiu. - Enviament per mail de l'arxiu de registre.	NO	- Programació per establir quan temps es pot fer servir Internet Explorer.	- Protecció per contrasenya. - Possibilitat d'excluir als superusaris de la màquina respecte el filtratge. - Activació o desactivació del filtre amb un sol botó.	- Windows 95, 98, NT, 2000, XP - Internet Explorer 5.01 o superior.
Optenet Web Filter PC 9.4.1 (OPTENET)[10]	- Llista negra (per categories temàtiques) - Llista blanca.	- Descàrrega de fitxers segons extensió. - P2P, Chat, Missatgeria Instantània,	- Webs visitades. - Webs bloquejades i motiu.	NO	SI	- Protecció per contrasenya.	- Windows 98, Me, 2000, XP. - Pentium o similar. - 64 MB RAM. - El programa actua com un proxy i la seva interfície és web.

Aplicació (autor)	Filtratge web	Bloquejos	Registre i Monitorització	Protecció privacitat	Limitació temps connexió	Altres	Requisits tècnics mínims / Observacions
		Email, grups de notícies i qualsevol port o rang de ports.					
SafeEyes 2006 (SafeBrowse.com)[11]	- Llista negra (per categories temàtiques) - Llista blanca. - Paraules prohibides.	- Finestres emergents. - Missategeria instantània. - Programes P2P, Jocs i qualsevol altra aplicació d'Internet.	- Webs visitades - Webs bloquejades i motiu. - Converses de chat i amb qui es parla.	NO	- Temps total de connexió a Internet per usuari. - Programació horària d'accés a Internet per usuari.	- Protecció per contrasenya. - Els informes d'activitat es poden guardar en local o en el servidor de l'empresa que ha fet el programa.	- Windows 98, Me, 2000, XP, MAC OSX Tiger 10.4 - Processador Intel o PowerPC. - 128 MB RAM. - 10 MB disc dur. - Resolució de pantalla 800x600. - Incorpora barra d'eines per Internet Explorer i per FireFox per tal d'accedir més fàcilment a les opcions del programa.
CyberPatrol 7.5 (CyberPatrol)[12]	- Llista negra (per categories temàtiques). - Llista blanca. - Paraules prohibides i paraules permeses.	-Estils de pàgina. - Missategeria instantània. - Programes P2P, Jocs i qualsevol altra aplicació d'Internet. - Descàrrega de fitxers segons extensió.	- Webs visitades - Webs bloquejades , motiu, temps de visita.	SI	- Temps total setmanal de connexió a Internet per perfil. - Programació horària d'accés a Internet per perfils.	- Protecció per contrasenya. - Encriptació dels registres d'activitat.	- Windows 98, Me, NT 4, 2000 Pro, XP. - Processador Pentium II o superior. - 64 MB RAM. - 30 MB disc dur. - Internet Explorer 4.0 SP2, Firefox 1.0, AOL 8.0, Netscape 6.0.

Nota:

[1] Internet Explorer: http://www.microsoft.com/windows/ie/ie6/
[2] Mozilla FireFox: http://www.mozilla.com/firefox/
[3] Opera: http://www.opera.com/
[4] NetNanny 5 http://www.netnanny.com/
[5] Control Kids http://www.controlkids.com/es/download.html
[6] FreeShield http://www.freeshield.com/
[7] InternetControllerQuattro http://controller.wazanet.net/quattro/
[8] FreeProxy http://www.handcraftedsoftware.org/index.php?page=5
[9] ChildWebGuardian http://www.childwebguardian.com/
[10] Optenet http://www.optenet.com
[11] SafeEyes http://www.safeeyes.com
[12] CyberPatrol http://www.cyberpatrol.com

1.2 Objectius del TFC

Des de fa us anys la competència en el mercat dels navegadors s'ha accentuat, i tot i que Internet Explorer continua tenint el lideratge, han sorgit alternatives de gran qualitat com els navegadors FireFox o Opera, per tant es lògic pensar que la tendència dels programes de filtratge web ha de ser la compatibilitat amb tots els navegadors lo que garantirà el seu ús pel màxim nombre d'usuaris com sigui possible.
És precissament la compatibilitat un dels objectius principals del projecte per tal de que sigui una solució oberta cosa que s'aconsegueix fent que l'aplicatiu es comporti com un proxy http ja que el protocol HTTP és un estàndard implementat per tots els navegadors.

Adicionalment, la novetat tecnològica és altre dels objectius del projecte: La tecnologia JAVA™ de Sun està fortament arrelada així que explorar les capacitats d'una altra tecnologia més nova com és el llenguatge de programació C# sota Microsoft® .NET Framework (la qual està disponible desde l'any 2002) i, relacionant-la amb la comunicació via xarxa, suposa un repte més interessant.

Finalment un objectiu més a destacar és aconseguir un aplicatiu de fàcil ús, de baix cost i petit tamany.

1.3 Enfocament i mètode seguit

L'estudi de com funcionen i quines característiques tenen les aplicacions de filtratge existents en el mercat ha estat imprescindible per determinar l'abast del projecte ja que això, encara que de forma indirecta, proporciona informació sobre les funcionalitat que poden necessitar els usuaris i de quines són les possibles tècniques per realitzar la implementació.

A causa de l'escassetat d'informació relacionada amb el camp del filtratge web usant la tecnologia de Microsoft i el llenguatge C#, el procés de cerca d'informació ha consistit en la recopil·lació de petites informacions útils i un procés d'integració i investigació constant per resoldre les casuístiques que han anat sorgint.

El procés d'implementació s'ha realitzat seguint una planificació per objectius de forma que cada vegada que s'aconseguia una fita es realitzaven proves. Això ha permès corregir errors i adaptar el codi en els casos en que el disseny havia de ser modificat continuament.

1.4 Planificació del Projecte

La planificació s'ha dividit en 6 fases:

- Recollida d'Informació
- Anàlisi i Disseny
- Implementació
- Proves i Correccions
- Documentació del Producte
- Memòria i Presentació

S'han produit desviacions a la planificació inicial a causa de les dificultats sorgides principalment, durant la implementació que han provocat canvis de disseny o obligat a cercar nova informació. En la següent taula es detalla tota la informació:

Tabla 1-2 Detall de la planificació

	FASE 1: Recollida d'Informació	FASE 2: Anàlisi i Disseny	FASE 3: Implementació	FASE 4: Proves i Correccions	FASE 5: Documentació del Producte	FASE 6: Memòria i Presentació
Període estimat inicialment	13 - 26 març del 2006 (2 setmanes)	27 març – 23 abril del 2006 (2 setmanes)	24 abril – 28 maig del 2006 (5 setmanes)	29 maig – 4 juny del 2006 (1 setmana)	5 –11 juny del 2006 (1 setmanes)	12 – 25 juny del 2006 (2 setmanes)
Període real	13 març - 20 abril	28 març – 5 maig	24 abril – 17 juny	26 abril – 17 juny	13 juny – 22 juny	17 juny – 22 juny (memòria) 22 juny – 24 juny
Descripció tasques	Recollida de la següent informació: -Característiques del protocol HTTP. -Programes de filtratge existents (com funcionen, característiques, etc). -Informació tècnica per realitzar l'anàlisi. -Disseny i implementació (possibles formes de funcionament de l'aplicació, Microsoft .NET Framework, llenguatge C#...)	Estudi de les possibilitats existents per desenvolupar l'aplicació: -Determinar l'abast del projecte. -Realització de l'anàlisi de l'aplicació segons les funcionalitats escollides. -Realització del disseny de l'aplicació segons l'anàlisi.	Implementació de l'aplicació (en primer lloc es realitzarà l'esquelet de l'aplicació i posteriorment s'aniran afegint la resta de funcionalitats) i realització de proves.	Realització de proves en situacions reals de funcionament , detecció i correcció de possibles problemes.	Realització de la documentació funcional del producte.	Realització de la memòria del projecte reunint tota la documentació realitzada fins el moment i afegint la informació requerida. Finalment, el·laboració d'una presentació clara i concisa del producte desenvolupat.
Objectius	-Aprofundir en el coneixement del protocol HTTP. -Aprofundir en el coneixement de Microsoft .NET i el C#. -Determinar la solució més adient per a la realització	-Escollir les funcionalitats que tindrà l'aplicació. -Determinar la solució més adient per a la realització del projecte.	-Implentació de l'aplicació segons el disseny. -Realització de proves per cada funcionalitat implementada. -Realització de proves una vegada s'ha finalitzat el desenvolupament.	-Assegurar la qualitat del producte final (millores, correccions, optimitzacions…)	-Obtenir una solució per a la instal·lació del producte. -Documentar les funcionalitats i possibles configuracions de l'aplicació.	-Reunir informació útil i necessària per la memòria. -Crear una presentació del producte desenvolupat.

	del projecte. -Obtenir un ventall de les possibles funcionalitats per implementar en el desenvolupament.					
Fites	-Document sobre característiques del protocol HTTP necessàries pel projecte. -Relació d'aplicacions de filtratge disponibles i característiques.	-Document d'anàlisi i disseny i justificació de les decisions preses.	-Creació de l'aplicació de filtratge de pàgines web. -Document de la implementació (solucions i tecnologies emprades, documentació de classes, etc)	-Ampliació en la documentació de proves. -Modifcació document de la implemtació en cas de correccions o modificacions	-Manual d'Usuari	-Memòria del projecte (lliurament 18 juny) -Presentació del producte. (lliurament 26 juny)
Motiu desviació	Escassetat d'informació que provoca retard en la decisió de quin seria el disseny adient. Necessària més cerca d'informació.	Es fa necessària més cerca d'informació	- Implementació pilot amb threading no satisfactòria (baix rendiment) -Cerca d'informació per implementar segons tècnica dels delegats asíncrons (threading basat en events). - La màquina de desenvolupament s'espatlla (3 dies perduts).	- Es realitzen proves a mesura que es desenvolupa. Les proves acaben més tard del previst a causa del d'una regressió en el codi la seva posterior correcció i en la investigació i correcció d'un problema amb els caràcters accentuats.	-Retard acumulat en fases anteriors.	-Retard acumulat en fases anteriors.

1.5 Productes obtinguts

El projecte té com resultat un executable de petit tamany (menys de 400Kb) que realitza les funcions de filtre web i 4 fitxers XML que contenen la configuració del programa.
A més s'inclou un manual explicant la instal·lació i la configuració i la documentació de les classes.

1.5.1 Umbra

L'aplicació de filtratge web Umbra[13] és un executable (de nom FiltroWeb.exe) per ser corregut sota Microsoft® .NET Framework 1.1. Està programat en C#, un llenguatge fortament orientat a objectes similar al Java i C++.

Les principals funcionalitats d'Umbra són:

- 4 modalitats de filtratge web
  - Bloqueig de totes les webs (restricció total)
  - Bloqueig per Llista Negra de Servidors i Llista de Paraules Prohibides (filtratge segons URL i contingut)
  - Permetre només a webs de la Llista Blanca (navegació segura)
  - Permetre tot (sense filtratge)
- Enregistrament de webs visitades i bloquejades
- Protecció de la configuració mitjançant clau encriptada.
- Configuració senzilla vía interfície gràfica de les opcions més comuns.
- Configuració manual vía fitxer de configuració de la resta d'opcions.

Nota:
[13] umbra-ae: del llatí, ombra, protecció, fantasma.

1.5.2 Fitxers de configuració

Junt amb l'executable s'inclouen 4 fitxers XML que són necessaris per executar l'aplicatiu. Aquests fitxers permeten guardar la configuració i determinen el filtratge que es realitzarà.

- Config.xml: conté la configuració d'Umbra.
- ListaBlancaServidores.xml: conté la llista de servidors que es consideren segurs.
- ListaNegraServidores.xml: conté la llista de servidors no segurs i que s'han de bloquejar.
- ListaPalabrasProhibidas.xml: conté la llista de paraules que provocaran que es bloquegi una web.

1.6 Resum dels continguts de la memòria

La memòria que es presenta a continuació inclou el procediment estàndar d'un desenvolupament de programari. En el capítol 2 hi trobem la recollida i documentació de requisits, un procés que pot ser considerarat en certa manera un preanàlisi.
En el següent capítol està englobada la fase d'anàlisi orientada a objectes així com informació tècnica relativa a la tecnologia que es vol emprar.
Finalment en el capítol 4 es mostra el disseny orientat a objectes, que es la conseqüència directa de la fase anterior, i que propociona una visió més acurada del mode de treball de l'aplicatiu desenvolupat.

2 Preanàlisi (Recollida i documentació de requisits)

2.1 Informació inicial

L'aplicació a desenvolupar filtrarà les pàgines web que es visitin per tal d'evitar que es visualitzin continguts no adequats d'acord a una sèrie de criteris configurables.
Aquests criteris podran ser configurats per un usuari administrador de l'aplicació i seran bàsicament una Llista Negra de Servidors on s'especificaran les URL que han de ser bloquejades, una Llista Blanca de Servidors que contindrà aquelles URL que es consideren segures i no s'han de bloquejar i una Llista de Paraules Prohibides a on s'especificaran totes aquelles paraules que provocaran que una web sigui bloquejada.

El filtratge de pàgines web constarà de 4 modalitats:

- Es bloquejarà l'accés segons la Llista Negra de Servidors i la Llista de Paraules Prohibides.
- Es permetrà l'accés a aquells servidors continguts a la Llista Blanca de Servidors.
- Bloqueig de totes les webs (incloent les que estiguin en els servidors de la Llista Blanca de Servidors)
- Bloqueig de totes les webs excepte les contingudes a la Llista Blanca de Servidors.

L'administrador també podrà guardar registres d'activitat per tal de guardar les webs visitades i bloquejades. Les opcions d'enregistrament seran:

- Per les URL visitades es guardarà: URL, usuari, data i hora, motiu d'haver-se permés (la URL era dins la Llista Blanca de Servidors, etc).
- Per les URL bloquejades es guardarà: URL, usuari, data i hora, motiu del bloqueig (la URL era dins la Llista Negra de Servidors, contenia paraules prohibides, etc).
- Les opcions de registre seran:
    - Enregistrar només les webs bloquejades.
    - Enregistrar les webs permeses i bloquejades
    - Desactivar l'enregistrament.

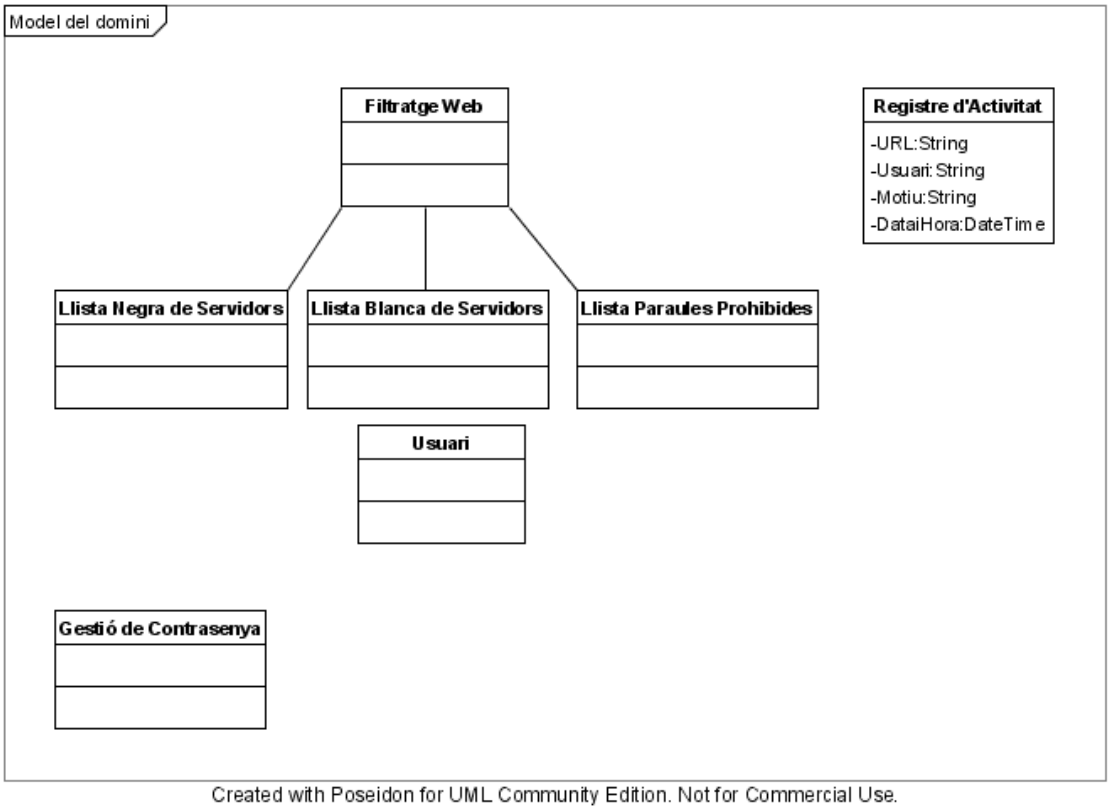
Les opcions configurables de l'aplicació estaran protegides per una contrasenya per tal de que només l'administrador pugui modificar-les, per tant serà necessària una opció per recordar la contrasenya en cas de pèrdua d'aquesta.

2.2 Model del domini

A primera vista s'identifquen els objectes 'Filtratge Web', 'Llista Negra de Servidors', 'Llista Blanca de Servidors', 'Llista Paraules Prohibides', 'Registre d'Activitat', 'Usuari' i 'Gestió de Contrasenya'. Evidentment les llistes estaran formades per objectes del tipus corresponent, és a dir, la 'Llista Negra de Servidors' i la 'Llista Blanca de Servidors' seran conjunts d'objectes 'Servidor' mentres que la 'Llista de Paraules Prohibides' estarà formada per objectes 'Paraula'.

Del registre d'activitat si que es poden intuïr els atributs i per aquest motiu el representem en el diagrama.

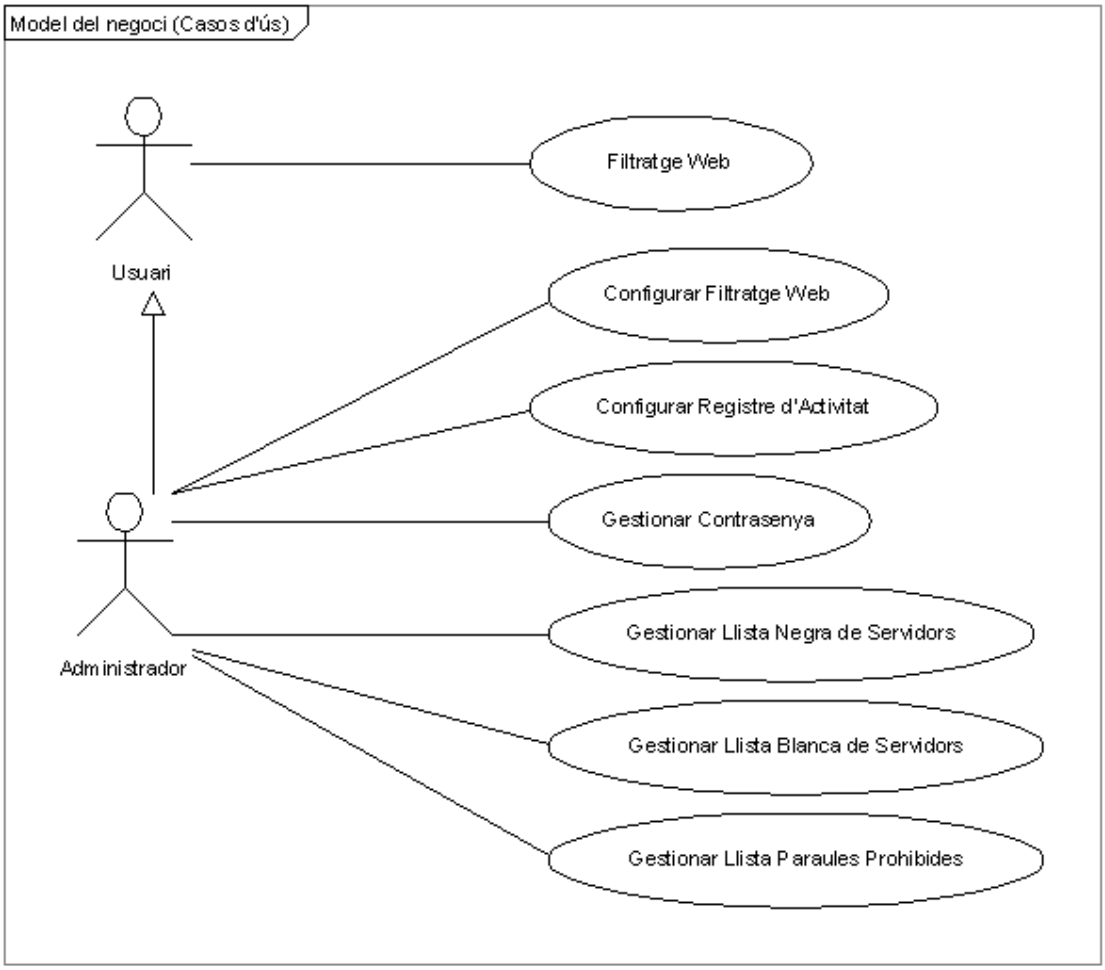
Figura 2-1 Model del domini



2.3 Model del negoci

El diagrama de casos d'ús del model del negoci és el següent:

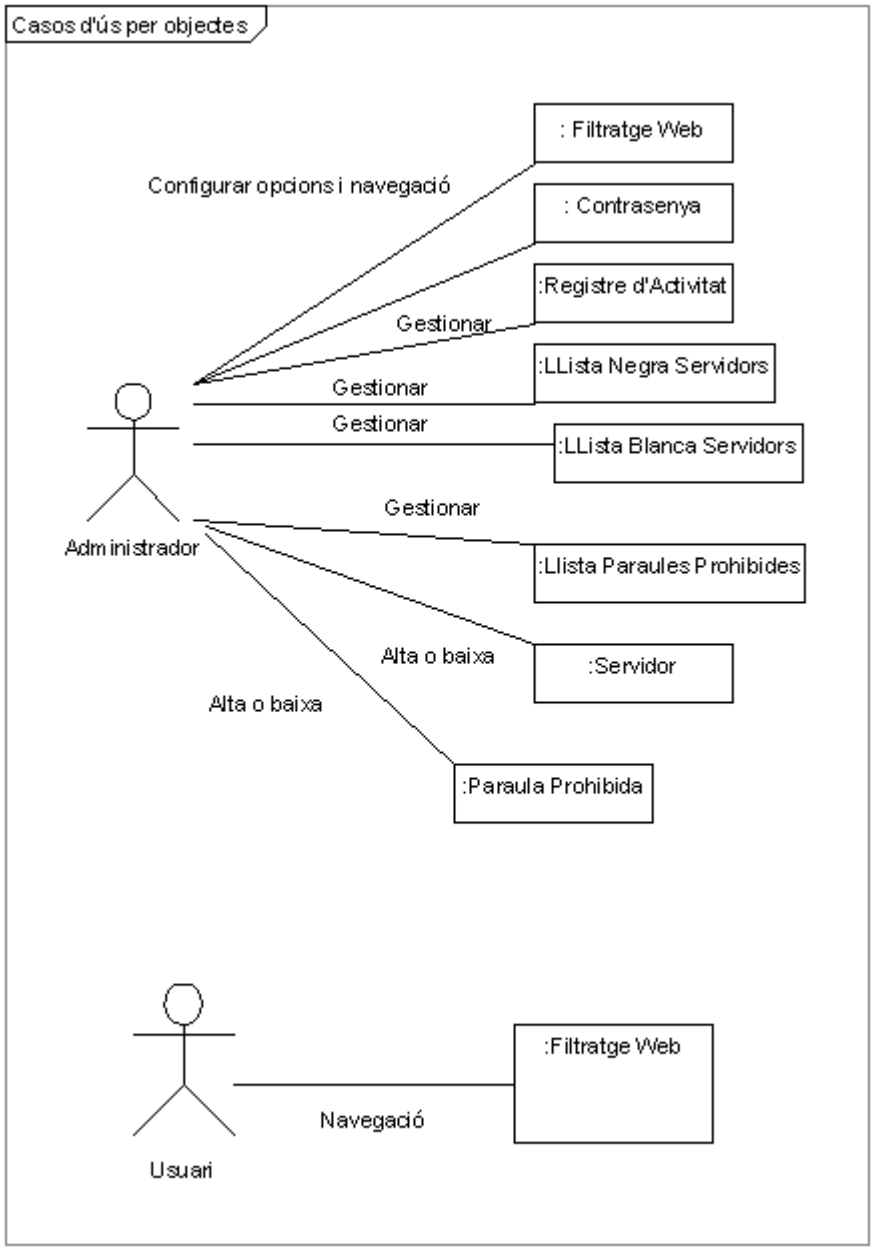
Figura 2-2 Model del negoci (Casos d'ús)



Com es pot veure els usuaris no administradors ("Usuari") només interaccionen amb el programa rebent el filtratge de les pàgines web mentre que l'Administrador té accés a totes les funcionalitats.

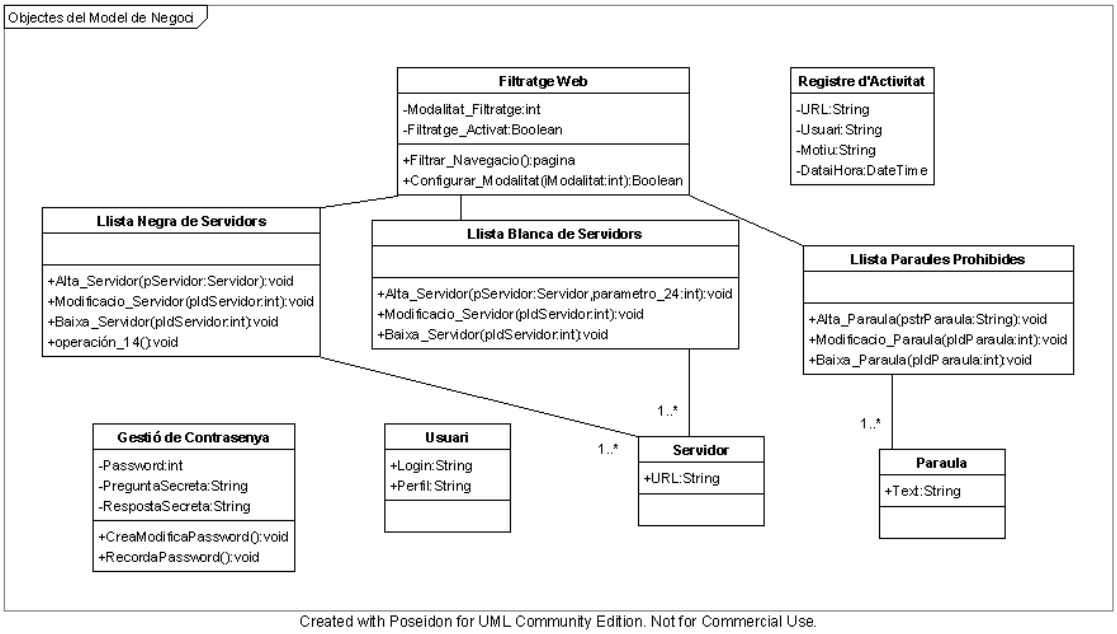
A partir dels casos d'ús es poden identificar els objectes que es fan servir en cadascun d'ells.

Figura 2-3 Casos d'ús per objectes



Partint de les dades anteriors obtenim una primera aproximació al diagrama d'objectes de l'aplicatiu:

Figura 2-4 Objectes del model de negoci



Com es pot veure el **Filtratge Web** fa servir la llistes de **servidors** i de **paraules prohibides**, de fet aquí és on constaran les modalitats de filtratge i a on s'indicarà si aquest està activat o no.
Per Servidor entendrem que és una URL ja que hom pot bloquejar, per exemple, totes les pàgines que tinguin "www.exemple.cat/recurs/" a la seva URL en comptes de bloquejar totes les del servidor especificant només "www.exemple.cat".

Cal dir que els atributs i mètodes que apareixen són provisionals i podrien canviar si l'anàlisi posterior així ho requereix.

2.4 Casos d'ús

A continuació es farà una relació dels casos d'ús que s'han identificat segons la situació plantejada.

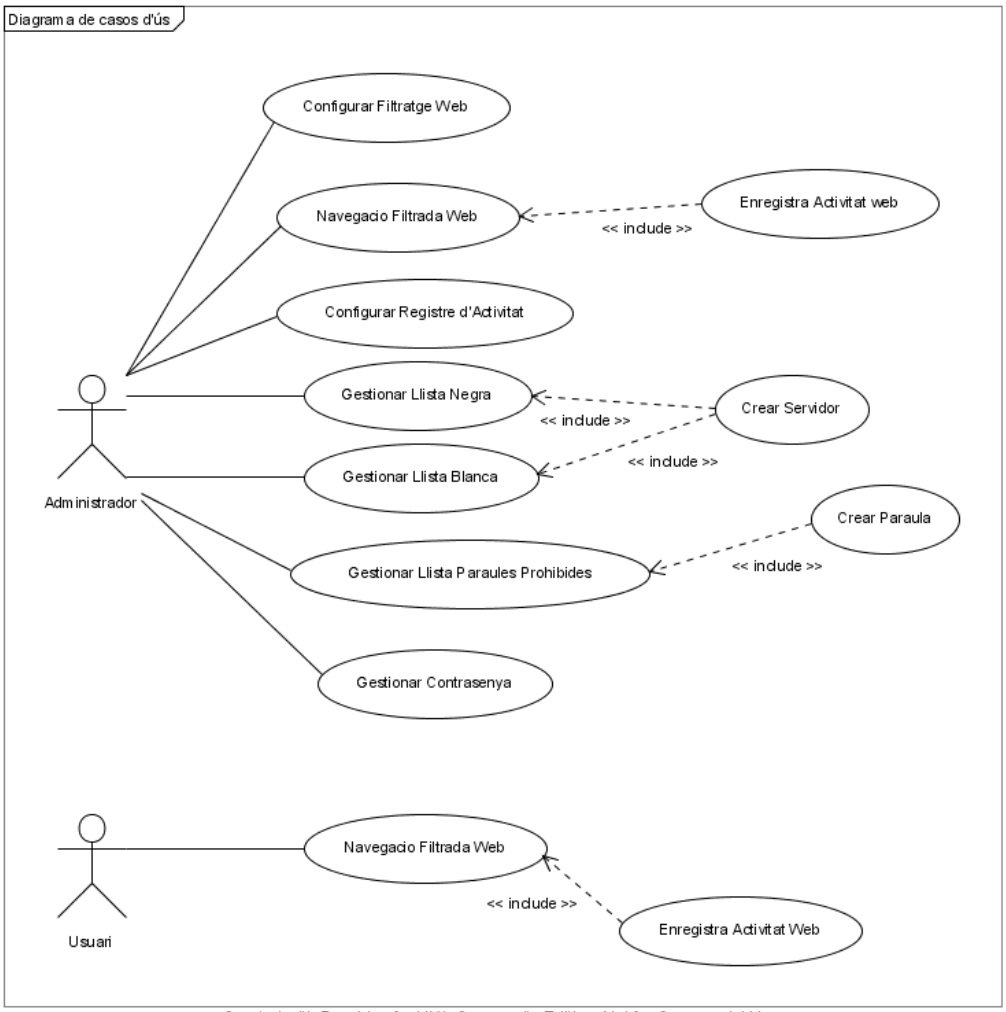
2.4.1 Actors

Els actors en seran dos: **usuari** i **administrador**. El primer d'ells és realment a qui va destinada la finalitat del programa: se li filtraran les pàgines web que visiti.
L'altre actor, l'**administrador** és qui configura el programa (tot i que també el programa també filtrarà les pàgines que visiti).

L'única diferència entre un ambdós actors és que un coneix la contrasenya d'accés a la configuració del programa.

2.4.2 Diagrama de casos d'ús

Figura 2-5 Diagrama de casos d'ús



2.4.3 Documentació textual dels casos d'ús

Cas d'ús número 1: "Configurar Filtratge Web"

Resum de la funcionalitat: configura la modalitat de filtratge a realitzar.
Paper dins l'operativa a implementar: és el cas d'ús principal per l'*administrador*.

Actors: **administrador**.
Casos d'ús relacionats: Gestionar Llista Negra, Gestionar Llista Blanca, Gestionar Paraules Prohibides, Gestionar Contrasenya, Navegacio Web Filtrada.
Precondició: S'ha introduit la contrasenya correcta.

Postcondició: El tipus de filtratge a aplicar queda configurat.

L'**administrador** introdueix la contrasenya correcta i a partir d'aqui pot configurar les opcions de filtratge: activar-lo o desactivar-lo, modificar la *Llista Negra de Servidors*, *Llista Blanca de Servidors* o la *Llista de Paraules Prohibides*. Per tant es podran afegir/modificar/esborrar *Servidors* o *Paraules.*

Cas d'ús número 2: "Navegacio Filtrada Web"

Resum de la funcionalitat: el programa filtres les webs visitades.
Paper dins l'operativa a implementar: és el cas d'ús principal per l'*administrador* i l'*usuari*. De fet és el cas d'ús més important de l'aplicació ja que és precissament la seva finalitat.

Actors: **administrador**, **usuari**.
Casos d'ús relacionats: Enregistra activitat web.
Precondició: el Filtratge Web ha d'estar activat.
Postcondició: les pàgines web són filtrades segons la configuració de filtratge.

L'aplicació filtra les pàgines web segons els criteris configurats.

Qüestions que cal aclarir:

• El filtratge s'aplica tant per l'**administrador** com per l'**usuari**? Resposta: sí; la funcionalitat de filtratge es permament sempre que estigui activa.

Cas d'ús número 3: "Configura Registre d'Activitat"

Resum de la funcionalitat: Selecciona el mode d'enregistrament de las webs visitades/bloquejades.
Paper dins l'operativa a implementar: es una funcionalitat que pot ser utilizada ocasionalment per l'***administrador*** (que és l'únic que la pot utilitzar).

Actors: **administrador**.
Casos d'ús relacionats: Gestionar Contrasenya, Navegacio Filtrada Web.
Precondició: L'*administrador* ha d'introduir la contrasenya correcta.
Postcondició: les pàgines web enregistrades segons els criteris configurats.

L'aplicació enregistra les pàgines web segons els criteris configurats: enregistrar només les webs bloquejades, les webs permeses i bloquejades o desactivar l'enregistrament.

Cas d'ús número 4: "Gestionar Llista Negra"

Resum de la funcionalitat: Consulta, afegeix, modifica o esborra *Servidors* de la *Llista Negra de Servidors.*
Paper dins l'operativa a implementar: és una funcionalitat que pot ser utilizada ocasionalment per l'***administrador*** (que és l'únic que la pot utilitzar) per tal d'incoporar/modificar *Servidors* que seran bloquejats.

Actors: **administrador**.
Casos d'ús relacionats: Gestionar Contrasenya, Configurar Filtratge Web, Navegacio Filtrada Web, Crear Servidor.

Precondició: L'*administrador* ha d'introduir la contrasenya correcta.
Postcondició: Si escau es realitzen els canvis pertinents sobre la *Llista Negra de Servidors*.

L'*administrador* pot consultar la *Llista Negra de Servidors* o modificar-la tot afegint o esborrant *Servidors* els quals seran bloquejats quan les opcions de filtratge indiquin que s'ha de filtrar usant aquesta llista.

Cas d'ús número 5: "Gestionar Llista Blanca"

Resum de la funcionalitat: Consulta, afegeix, modifica o esborra *Servidors* de la *Llista Blanca de Servidors.*
Paper dins l'operativa a implementar: és una funcionalitat que pot ser utilizada ocasionalment per l'***administrador*** (que és l'únic que la pot utilitzar) per tal d'incoporar/modificar *Servidors* que no seran bloquejats.

Actors: **administrador**.
Casos d'ús relacionats: Gestionar Contrasenya, Configurar Filtratge Web, Navegacio Filtrada Web, Crear Servidor.
Precondició: L'*administrador* ha d'introduir la contrasenya correcta.
Postcondició: Si escau es realitzen els canvis pertinents sobre la *Llista Blanca de Servidors*.

L'*administrador* pot consultar la *Llista Blanca de Servidors* o modificar-la tot afegint o esborrant *Servidors* els quals no seran bloquejats quan les opcions de filtratge indiquin que no s'han de filtrar aquells que hi constin dins d'ella..

Cas d'ús número 6: "Gestionar Paraules Prohibides"

Resum de la funcionalitat: Consulta, afegeix, modifica o esborra *Paraules* de la *Llista de Paraules Prohibides.*
Paper dins l'operativa a implementar: és una funcionalitat que pot ser utilizada ocasionalment per l'***administrador*** (que és l'únic que la pot utilitzar) per tal d'incoporar/modificar *Paraules* que provocaran que es bloquegin aquelles webs que les continguin.

Actors: **administrador**.
Casos d'ús relacionats: Gestionar Contrasenya, Configurar Filtratge Web, Navegacio Filtrada Web, Crear Paraula.
Precondició: L'*administrador* ha d'introduir la contrasenya correcta.
Postcondició: Si escau es realitzen els canvis pertinents sobre la *Llista de Paraules Prohibides*.

L'*administrador* pot consultar la *Llista de Paraules Prohibides* o modificar-la tot afegint o esborrant *Paraules* les quals provoquen que aquelles webs que les continguin siguin bloquejades quan les opcions de filtratge indiquin que s'ha de filtrar fent servir aquesta llista.

Cas d'ús número 7: "Gestionar Contrasenya"

Resum de la funcionalitat: Permet establir la contrasenya administrativa, la qual permet modificar les totes les opcions del programa.

Paper dins l'operativa a implementar: és una funcionalitat que pot ser utilizada ocasionalment per l'***administrador*** (que és l'únic que la pot utilitzar) per tal de modificar la contrasenya administrativa.

Actors: **administrador**.
Casos d'ús relacionats: Configurar Filtratge Web, Gestionar Llista Paraules Prohibides, Gestionar Llista Negra de Servidors, Gestionar Llista Blanca de Servidors, Configurar Registre d'Activitat.
Precondició: L'*administrador* ha d'introduir la contrasenya correcta per poder gestionar la contrasenya actual.
Postcondició: Si escau es realitzen els canvis pertinents sobre la contrasenya.

L'*administrador* pot modificar la contrasenya que permetrà la configuració de totes les opcions del programa.

Qüestions que cal aclarir:

• Quan s'acaba d'instal·lar el programa, com es pot gestionar la contrasenya? Resposta: A la persona que accedeix a les opcions de configuració del programa se li demana que estableixi una contrasenya i per tant es convertirà en l'**administrador**.

2.5 Requisits de la interficie d'usuari

La interficie d'usuari estarà orientada totalment per ser utilitzada per l'actor **administrador** ja que la seva finalitat serà la de configurar l'aplicació i consultar l'activitat web enregistrada (registres d'activitat). Es pretèn que sigui senzilla i funcional.

2.5.1 Perfils d'usuari

L'actor **usuari** no tindrá accés a la interficie d'usuari ja que caldrà introduir una contrasenya i aquesta es coneguda només per l'actor **administrador**.
L'administrador serà un usuari amb certa experiència amb ordinadors per tal de poder configurar l'aplicació correctament.

2.5.2 Requisits d'usabilitat

Tenint en compte que el nombre de funcionalitats accessibles per l'usuari no és gaire gran l'objectiu es que l'**administrador** sigui capaç de configurar i entendre el funcionament en menys d'una hora.

3 Anàlisi orientat a objectes

3.1 Modificacions respecte el Preanàlisi

S'ha cregut oportú convenient diferenciar clarament les opcions de filtratge del programa així com l'existència d'una modalitat que permeti passar tot el tràfic per cubrir la necessitat de que potser en algún moment convindria no realitzar cap filtratge.

Per tant, el filtratge de pàgines web constarà de 4 modalitats independents:

- Bloqueig de totes les webs (restricció total)
- Bloqueig per Llista Negra de Servidors i Llista de Paraules Prohibides (filtratge segons URL i contingut)
- Permetre només a webs de la Llista Blanca (navegació segura)
- Permetre tot (sense filtratge)

Així mateix amb l'objectiu de simplificar el sistema d'enregistrament de l'activitat del programa s'ha optat per guardar totes les webs en el mateix fitxer, tant si han estat bloquejades com si han estat permeses. Conseqüentment les opcions d'enregistrament seran les següents:

- Activar l'enregistrament (s'enregistraran les webs permeses i bloquejades)
- Desactivar l'enregistrament (no s'enregistra cap web)

En quant a la gestió de la contrasenya es prescindirà de la opció de recordatori de clau per simplificar-ne al màxim la interfície d'usuari.

3.2 Solucions tecnològiques i eines

Per a la implementació del programari s'utilitzarà la **tecnologia Microsoft® .NET Framework** (versió 1.1) amb el **llenguatge orientat a objectes C#** i l'ús de **XML** pels fitxers de configuració i registre.
El desenvolupament es farà amb **Visual Studio .NET 2003** una eina que integra totes les funcionalitats necessàries per desenvolupar aplicacions visuals.

Contràriament a la majoria del programes existents en el mercat que només funcionen amb Internet Explorer, l'aplicatiu a desenvolupar funcionarà per qualsevol navegador ja que el disseny es basarà en l'ús d'un **proxy HTTP 1.0** (sense cache).

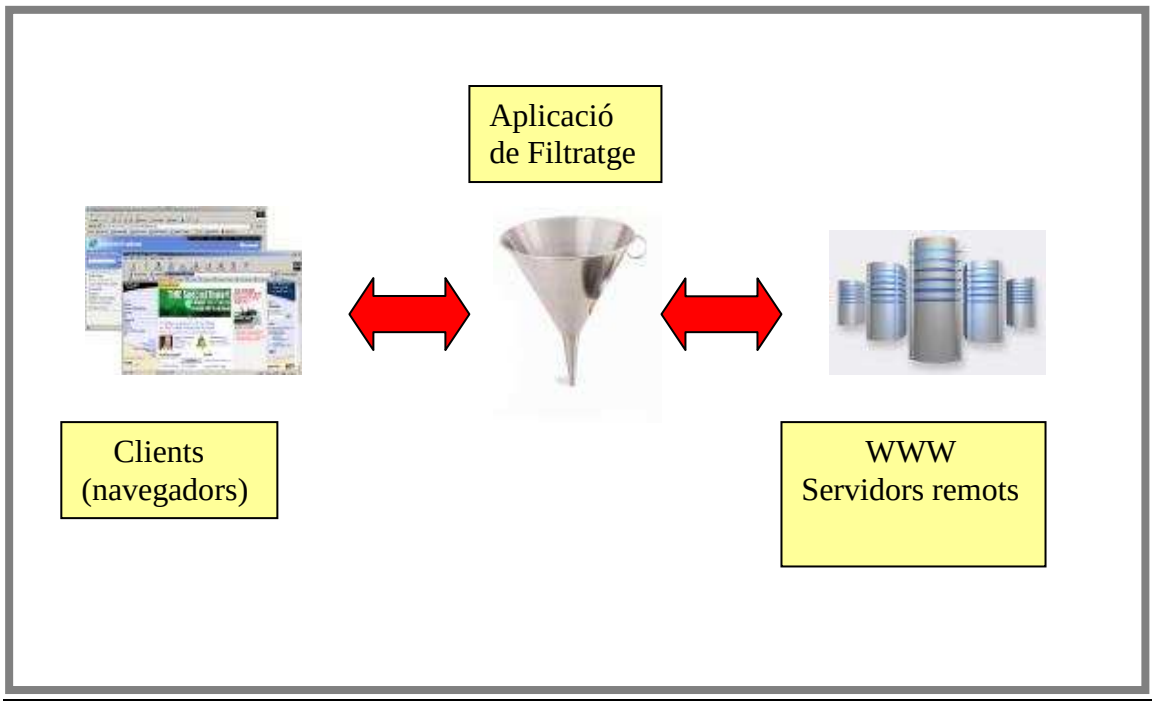
S'ha optat per aquest disseny ja que el proxy es situa entre els clients (navegadors) i els servidors web remots proporcionant els següents avantatges:

a) Garanteix la compatibilitat amb qualsevol navegador ja que usa el **protocol HTTP** per l'intercanvi de dades entre el proxy i el navegador.
b) Permetrà filtrar els continguts de les pàgines webs abans que arribin al client afegint la lògica adequada.

El proxy acceptarà múltiples connexions dels clients de forma que es poden fer connexions simultànies a diferents hosts remots en un mateix instant. Això serà possible mitjançant l'ús de **programació asincrònica amb C#,** que proporciona major modularitat (i per tant major nombre de funcions) en el codi a diferència dels *threads* (fils) en estar basada en un sistema similars als events.

El mòdul de filtratge requerirà extreure el contingut de les pàgines quan filtri buscant paraules prohibides dins el seu contigut. Com que les pàgines **HTML** contenen etiquetes caldrà distingir entre aquestes i el contigut real de la pàgina cosa que s'aconseguirà fent servir les ***Regular Expressions*** (expressions regulars) de **.NET Framework** les quals permeten detectar qualsevol patró de text que es defineixi.

Figura 3-1 Esquematització de la forma de treball



3.3 Protocol HTTP

3.3.1 Introducció al protocol HTTP

L'HTTP (*Hypertext Transfer Protocol*) és un protocol simple, orientat a conexió (client-servidor) i sense estat (quan el servidor respon a un client no guarda l'estat anterior en que es trobava el client). Es tracta d'un protocol orientat a connexió perquè fa servir el protocol de comunicacions TCP (*Transport Control Protocol*) de mode conectat que estableix un canal de comunicacions d'extrem a extrem (client-servidor)

Existeix una versió d'HTTP anomenada HTTPS (*Hypertext Transfer Protocol over Secure Socket Layer*) que utilitza el protocol de seguretat SSL (*Secure Socket Layer*).

El funcionament del protocol HTTP es basa en que un client estableix una connexió TCP cap a un servidor, en concret al port HTTP (normalment el 80) i envia una comanda HTTP de petició de recurs amb unes capçaleres informatives. El servidor respon per la mateixa connexió enviant la informació demanada i capçaleres informatives.

El protocol HTTP es troba actualment en la versió 1.1 i defineix com s'envien els paramàtres entre les pàgines web, l'existència de servidors de cache, etc. El protocol HTTP 1.0 encara és vigent.

Les directives o mètodes de les sol·licituds del protocol HTTP 1.1 són:

GET: sol·licitud de recurs.
POST: sol·licitud de recurs passant paràmetres.
HEAD: sol·licitud de dades sobre el recurs. Similar al GET excepte en que només es rep la resposta i no les dades de la pàgina.
PUT: Creació o enviament de recurs.
DELETE: eliminació de recurs.
TRACE: retorna a l'origen la sol·licitud tal com s'ha rebut en el receptor (amb la finalitat de depurar errors).
OPTIONS: comprovació de les capacitats del servidor.
CONNECT: reservada per l'ús en servidors intermitjos capaços de funcionar com tunels.

Cal destacar que el protocol HTTP 1.0 només especifica les 3 primeres: GET, POST i HEAD.

D'aquestes directives es detallen les més importants de cara a les sol·licituds dels navegadors: GET i POST.

Tant les peticions com les respostes tenen el següent format:

Una linia inicial
Cap o més línies de capçaleres.
Linia en blanc (per exemple amb un només un CR i LF –*Carriage Return* i *Line Feed*-).
Un cos del missatge opcional (fitxer, dades de consulta o de sortida)

Altra forma de representar-ho seria aquesta:

<linia inicial, diferent segons si es petició o resposta>

Capçalera1: valor1
Capçalera2: valor2
Capçalera3: valor3
[linia en blanc]
<cos del missatge opcional: contingut de fitxer, dades de consulta o de sortida. Pot tindre moltes linies o inclús dades binaries>
Les linies inicials i capçalares han d'acabar en CR i LF, encara que ho poden fer simplement amb LF (en ASCII CR i LF són els valors 13 i 10).

3.3.2 Peticions HTTP: GET i POST

Les peticions HTTP es realitzen mitjançant dos mètodes: GET i POST.
En el mètode GET s'envien els paràmetres codificats en la URL, en canvi en el mètode POST aquests s'envien dins del cos de la petició HTTP.

Exemple de petició GET:

```
GET /index.php HTTP/1.1
Host: www.exemple.cat
User-Agent: Mozilla/4.5 [en]
Accept: image/gif, image/jpeg, text/html
Accept-language: en
Accept-Charset: iso-8859-1
```

La petició está formada per:

A. Linia de petició

La linia de petició conté dades sobre el recurs que es sol·licita al servidor i està formada per:
1. Mètode: nom del método de HTTP (GET, POST, etc.).
2. Identificador de recurs: URL (*Uniform Resource Locator*)
3. Versión de protocolo: versión del protocolo sol·licitat per la resposta.

B. Capçalera de petició

Conté informació adicional per ajudar al servidor (o als servidors *proxy i cache*) a processar correctament la petició.
El format d'aquesta informació es:

Identificador: valor

Els identificador més importants són:

Host: nom del servidor sol·licitat (aquest identificador no és obligatori per HTTP 1.0)
User-Agent: nom del navegador que accedeix al recurs.
Accept: formats que accepta el client.
Accept-Language: idiomes preferits pel client.

C. Paràmetres de petició

Una petició HTTP pot incloure paràmetres (com els que s'envien quan omplim i enviem un formulari a una web) i aquests es poden passar dins la mateixa cadena de petició (codificats a la URL) o bé dins la mateixa petició.

En el primer cas, els paramàtres s'afegeixen darrera de la URL precedits del caràcter **?** i separats pel caràcter **&**. Els espais es substitueixen pel signe més (+) i qualsevol altre caràcter especial es representa amb la cadena %xx on xx és el codi ASCII en hexadecimal. Aquesta forma de codificar s'anomena **codificació URL**.

Per exemple si a la URL http://www.exemple.cat/index.php se li passessin els paràmetres nomprotocol y Alta amb els valors "Protocol http" i "1" respectivament seria la següent:

```
http://www.exemple.cat/index.php?nomprotocol=Protocol+http&Alta=1
```

Aquesta sol·licitud per part d'un navegador genera una petició HTTP del tipus GET i quedaria així:

```
GET /index.php?nomprotocol=Protocol+http&Alta=1 HTTP/1.0
Host: www.exemple.cat
User-Agent: Mozilla/4.5 [en]
Accept: image/gif, image/jpeg, text/html
Accept-language: en
Accept-Charset: iso-8859-1
```

En el cas de que els paràmetres es passin dins de la petició aquesta seria del tipus POST i tindria la forma següent:

```
POST /index.php HTTP/1.0
Host: www.exemple.cat
User-Agent: Mozilla/4.5 [en]
Accept: image/gif, image/jpeg, text/html
Accept-language: en
Accept-Charset: iso-8859-1

nomprotocol=Protocol+http&Alta=1
```

Com es veu a l'exemple anterior les peticions POST i GET són diferents encara que una petició POST també pot tenir paràmetres dins la linia de petició.
Els paràmetres que es passen dins el cos de la petició estàn codifcats segons la codificació URL però també ho poden estar en la **codificació multipart** la qual es basa en el format MIME (*Multipurpose Internet Mail Extensions*) i que també serveix per enviar fitxers al servidor.

L'exemple anterior en codificació multiparte seria el següent:

```
POST /index.php HTTP/1.0
Host: www.exemple.cat
User-Agent: Mozilla/4.5 [en]
Accept: image/gif, image/jpeg, text/html
Accept-language: en
Accept-Charset: iso-8859-1
Content-Type: multipart/form-data,
delimiter="----ALEATORI----"

----ALEATORI----
Content-Disposition: form-data; name="nomprotocol"
Protocol http
----ALEATORI----
Content-Disposition: form-data; name="Alta"
1

----ALEATORI------
```

3.3.3 Respostes HTTP

Les respostes del protocol HTTP són molt semblants a les peticions.
Una resposta estàndard podria ser similar a aquesta:

```
HTTP/1.1 200 OK
Date: Mon, 23 Mar 2006 17:21:24 GMT
Server: Apache/2.0.53 (Win32) PHP/5.0.3
Last-Modified: Fri, 13 May 2005 20:52:03 GMT
Accept-Ranges: bytes
Content-Length: 245
Connection: close
Content-Type: text/html

<html>
...

</html>
```

A la primera linia s'indica la versión del protocol amb el qual s'envia la pàgina web seguida d'un codi d'estat i una frase de retorn.
El codi de retorn pot tenir aquests valors (xx són números):

1xx: Indica un missatge informatiu.
2xx: Indica èxit d'algun tipus de la petició.
3xx: Redirecció del client a una altra URL.
4xx: Error de cliente. No es pot processar la petició.
5xx: Error de servidor.

La frase de retorn depèn de la implementació del servidor web i només serveix per aclarir el codi d'estat. Es pot veure la llista completa dels codis d'estat de l'especificació del protocol HTTP 1.1 a http://www.w3.org/Protocols/rfc2616/rfc2616-sec6.html#sec6.1.1

Les linies posteriors són camps de control amb el mateix format que les capçaleres de la petició que ens informen del contingut (data de creació, servidor i versió d'aquest, format del contigut, etc).

Nota:

- Especificació del protocol HTTP 1.0 http://www.w3.org/Protocols/rfc1945/rfc1945
- Especificació del protocol HTTP 1.1 http://www.w3.org/Protocols/rfc2616/rfc2616.html

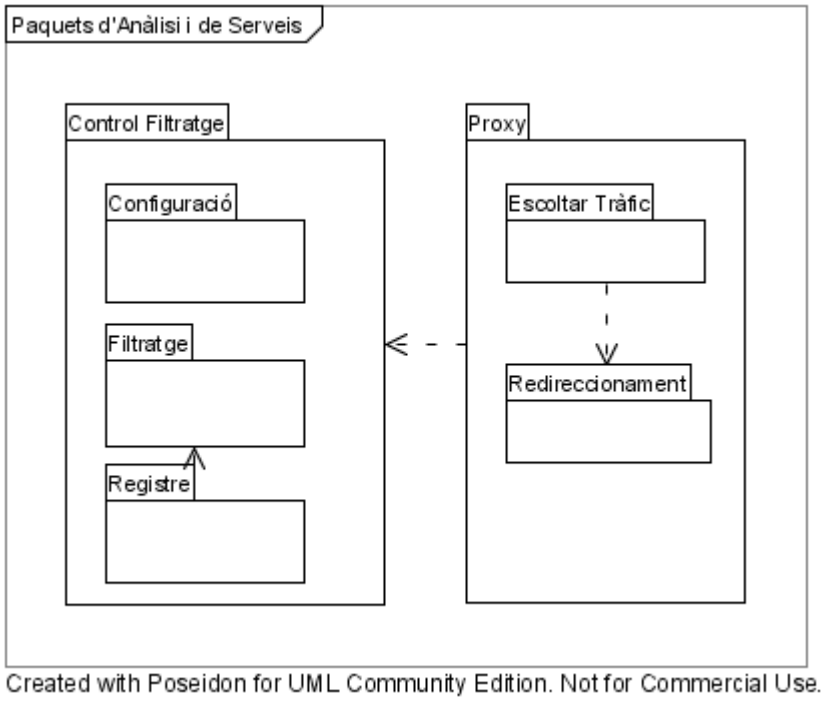
3.4 Requisits tècnics mínims

- Windows 2000/XP.
- Microsoft .NET Framework 1.1. instal·lat.
- 512 MB lliure de memòria RAM.
- Connexió a Internet de banda ampla.

3.5 Paquets d'anàlisi i serveis

Es poden diferenciar 2 paquets d'anàlisi **Control Filtratge** i **Proxy**. El primer d'ells inclou paquets de serveis per la **configuració** del filtre web, **registre** d'activitat i les que són específiques per determinar el **filtratge** de les webs.
El segon paquet implementa les funcionalitats d'un proxy HTTP de forma que els clients (navegadors) es conectaran a ell i aquest serà capaç de filtrar el tràfic amb la intervenció del paquet **Control Filtratge** (per aquest motiu tenen una relació de dependència). Com tot proxy consta d'una funcionalitat que escolta el tràfic entrant (representada pel paquet de servei **Escoltar Tràfic)** i altra que realitza el **redireccionament** de les conexions entre els clients i els servidors remots.

Figura 3-2 Paquets d'anàlisi i de serveis



3.6 Identificació de les classes de les entitats

A partir dels casos d'ús s'identifiquem les classes a implementar. Així tenim **Cliente**, **ClienteHTTP**, **EscuchadorHTTP** que integren el proxy; **RegistroActividad**, **GestionContrasenya**, **FiltradoWeb**, **Servidor**, **ListaServidores** i **ListaPalabrasProhibidas** que són les que controlen l'aplicació (configuració, registre d'activitat i el filtratge). La classe **Comun** conté mètodes de suport i constants.

3.6.1 Relacions entre entitats

E mòdul principal de l'aplicació és **FiltradoWeb** i hereta els mètodes i atributs de les classes de **RegistroActividad** per herència directa i de **GestionContrasenya** per herència indirecta. El motiu de que **FiltradoWeb** hereti de **GestionContrasenya** indirectament és deu a que C# no permet l'herència múltiple.

FiltradoWeb també és l'encarregat de controlar el filtratge i necessita de les dades contingudes a **ListaPalabrasProhibidas** i **ListaServidores**. De fet, aquesta última classe serà instanciada dues vegades per contindre la Llista Blanca de Servidors i la Llista Negra de Servidors.

El control de les connexions HTTP es centralitza en **EscuchadorHTTP** que crea nous clients a mesura que es reben sol·licituds HTTP dels clients. Cadasuna d'aquests sol·licituds provoca la creació d'un objecte **ClienteHTTP** que s'encarregarà de controlar el fluxe entre els clients i els servidors remots.

3.7 Casos d'ús

No hi ha canvis d'importància en el casos d'ús respecte la recollida i documentació de requisits. L'únic canvi és la desaparició de la funcionalitat Crear Paraula inclosa dins de Gestiona Lista Blanca i Gestiona Lista Negra a causa de que la classe Paraula ha estat eliminada en aquesta fase com s'ha esmentat anteriorment.

Figura 3-3 Diagrama de casos d'ús de l'anàlisi

3.7.1 Actors

L'aplicatiu té dos actors que identifiquem com **Usuari** i **Administrador**. El primer d'ells és realment a qui va destinada la finalitat del programa: se li filtraran les pàgines web que visiti.
L'altre actor, l'**Administrador** és qui configura el programa (tot i que també el programa també filtrarà les pàgines que visiti).

L'única diferència entre un ambdós actors és que un coneix la contrasenya d'accés a la configuració del programa i l'altre no, per tant a nivell d'implementació els dos actors son gairebé el mateix.

3.7.2 Identificació de les classes de frontera, de control i operacions

En aquest cas les classes de frontera i de control no són rellevants perquè l'aplicatiu es desenvolupa amb Visual Studio .NET 2003 que és un entorn integrat de desenvolupament on la creació y control dels menús es generada automáticament.

Cas d'ús número 1: "Configurar Filtratge Web"

Resum de la funcionalitat: configura la modalitat de filtratge a realitzar.
Paper dins l'operativa a implementar: és el cas d'ús principal per l'*administrador*.

Actors: **administrador**.
Operacions principals:

1. Canvia tipus de filtratge i canvia estat filtratge (classe FiltradoWeb)

Cas d'ús número 2: "Navegacio Filtrada Web"

Actors: **administrador**, **usuari**.
Operacions principals (en ordre d'execució):

1. Crea ClienteHTTP.
2. IniciarComunicacion (classe ClienteHTTP).
3. EsValida_SolicitudHTTP (classe Comun).
4. Procesar_Solicitud (classe ClienteHTTP).
5. IniciaRedirecciomiento (classe ClienteHTTP).
6. Bloquear_Web (classe ClienteHTTP).
7. Precarga_PaginaWeb (classe ClienteHTTP).
8. Existe_Direccion_Web (classe ClienteHTTP).
9. Existe_Palabra_Prohibida (classe ClienteHTTP).

Cas d'ús número 3: "Configura Registre d'Activitat"

Actors: **administrador**.
Operacions principals:

1. Canvia estat de l'enregistrament (classe FiltradoWeb i RegistroActividad) → Grabar_Fichero_Registro (classe Comun) → Borrar_RegistroActividad (classe FiltradoWeb i RegistroActividad)

Cas d'ús número 4: "Gestionar Llista Negra"

Actors: **administrador**.
Operacions principals:

1. Crear Servidor → Establece_Valores (classe Servidor) → Llenar_ListView (classe Comun i FiltradoWeb).
2. Borrar_Servidor_ListaNegra (classe FiltradoWeb)

Cas d'ús número 5: "Gestionar Llista Blanca"

Actors: **administrador**.
Operacions principals:

1. Crear Servidor → Establece_Valores (classe Servidor) → Llenar_ListView (classe Comun i FiltradoWeb)
2. Borrar_Servidor_ListaBlanca (classe FiltradoWeb)

Cas d'ús número 6: "Gestionar Paraules Prohibides"

Actors: **administrador**.
Operacions principals:

1. Insertar_Palabra_ListaPalabrasProhibidas (classe FiltradoWeb) → Llenar_ListView (classe Comun).
2. Borrar_Palabra_ListaPalabrasProhibidas (classe FiltradoWeb) → Llenar_ListView (classe Comun).

Cas d'ús número 7: "Gestionar Contrasenya"

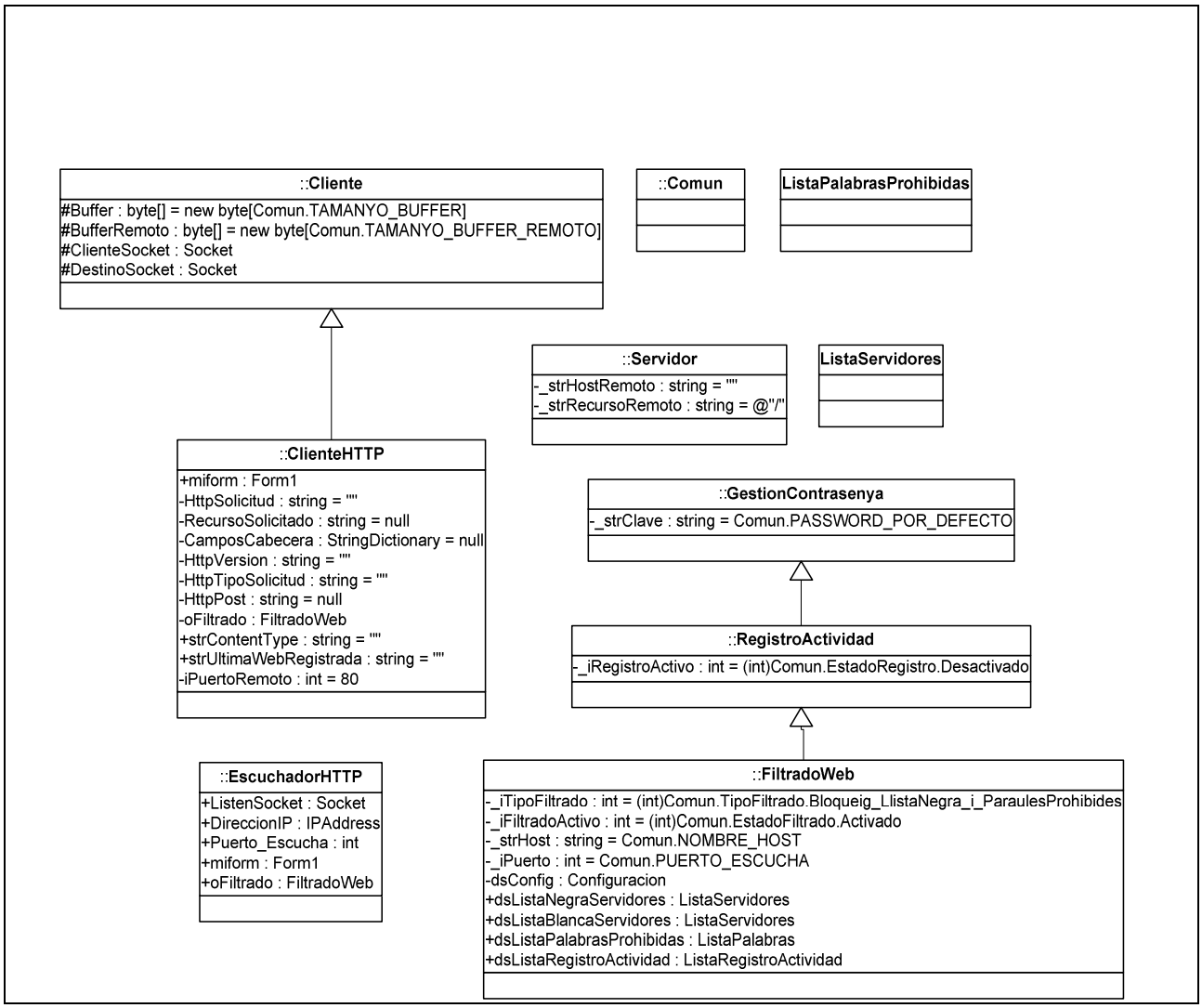
Actors: **administrador**.
Operacions principals:

1. Encriptar_Clave (classe GestionContrasenya) → assignar nova contrasenya (classe FiltradoWeb).

3.8 Diagrama estàtic d'anàlisi

El diagrama de les relacions entre classes és el següent

Figura 3-4 Diagrama estàtic d'anàlisi



3.9 Interfície gràfica

La interficie d'usuari estarà orientada totalment per ser utilitzada per l'actor **administrador** ja que la seva finalitat serà la de configurar l'aplicació i consultar l'activitat web enregistrada (registres d'activitat). Serà senzilla i funcional.

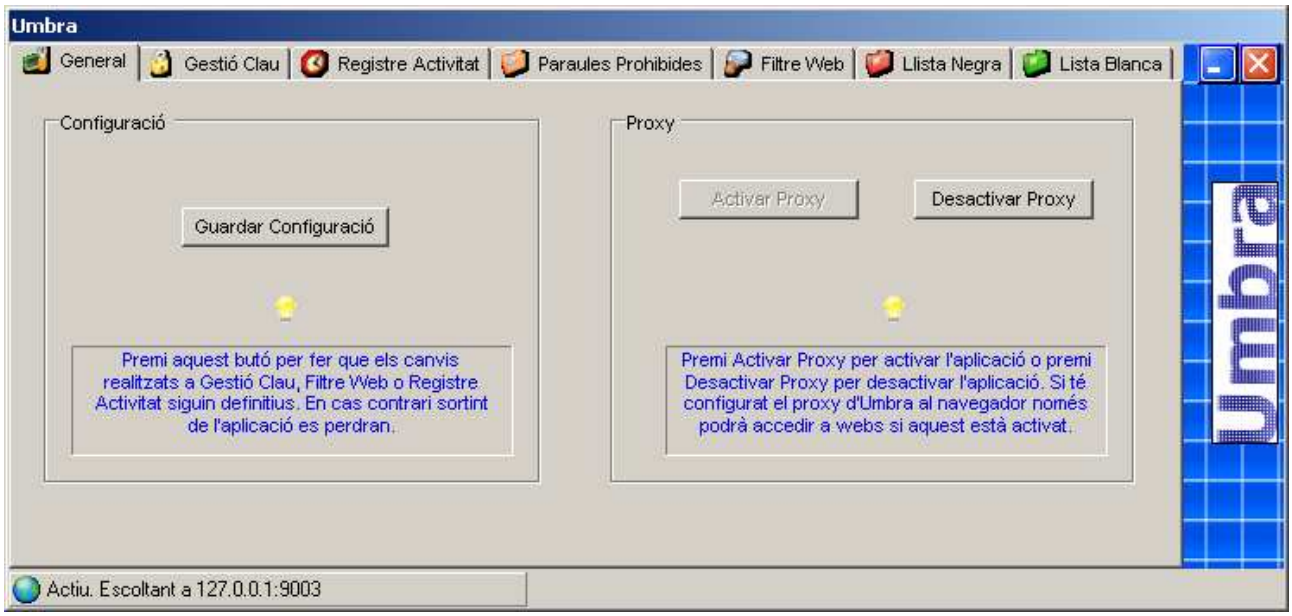
3.9.1 Pantalla de control d'accés

Figura 3-5 Pantalla de control d'accés

3.9.2 Pantalla principal de configuració de l'aplicació: guardar configuració i activar/desactivar proxy

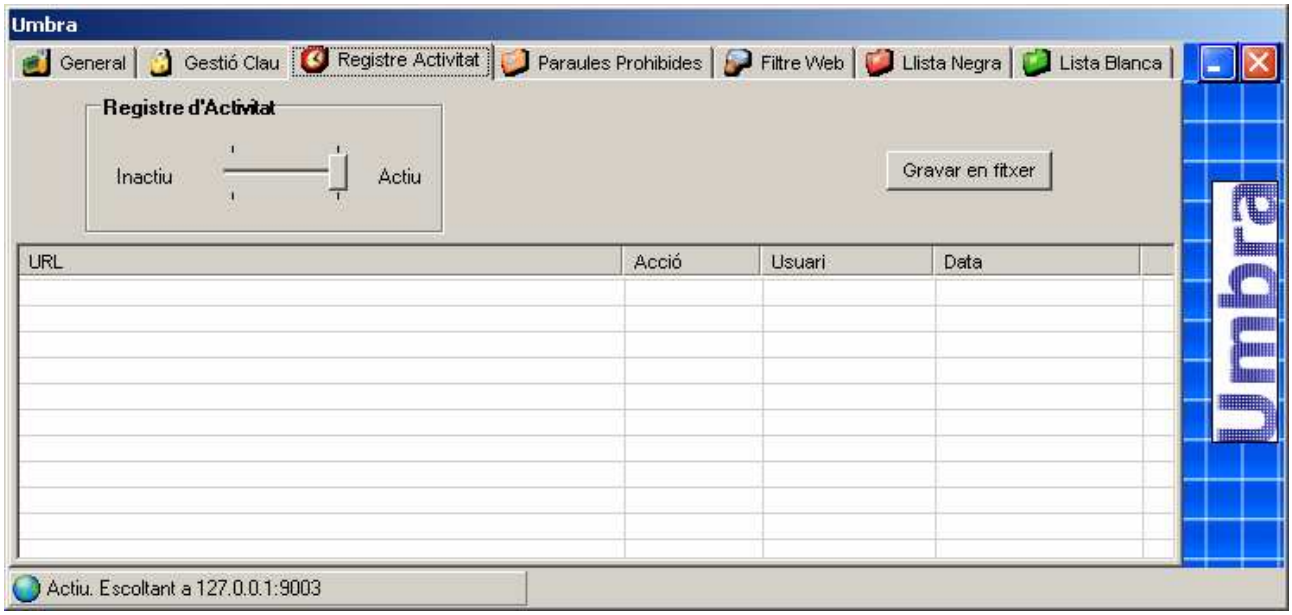
Aquesta pantalla permet guardar en disc la configuració del registre d'activitat, del filtratge i la clau així com activar o desactivar l'escoltament de connexions HTTP (activació/desactivació del proxy).

Figura 3-6 Pantalla principal de configuració



3.9.3 Pantalla de configuració del registre d'activitat

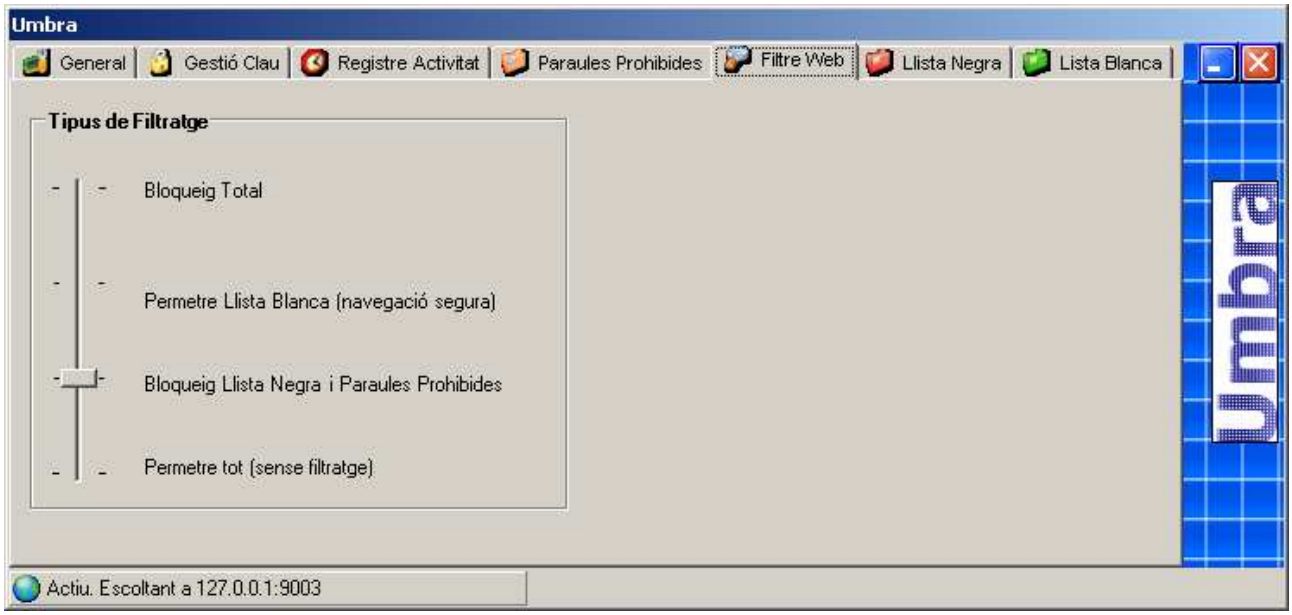
Figura 3-7 Pantalla de configuració del registre d'activitat



3.9.4 Pantalla de configuració del filtratge web

Des de aquesta pantalla es pot configurar el tipus de filtratge.

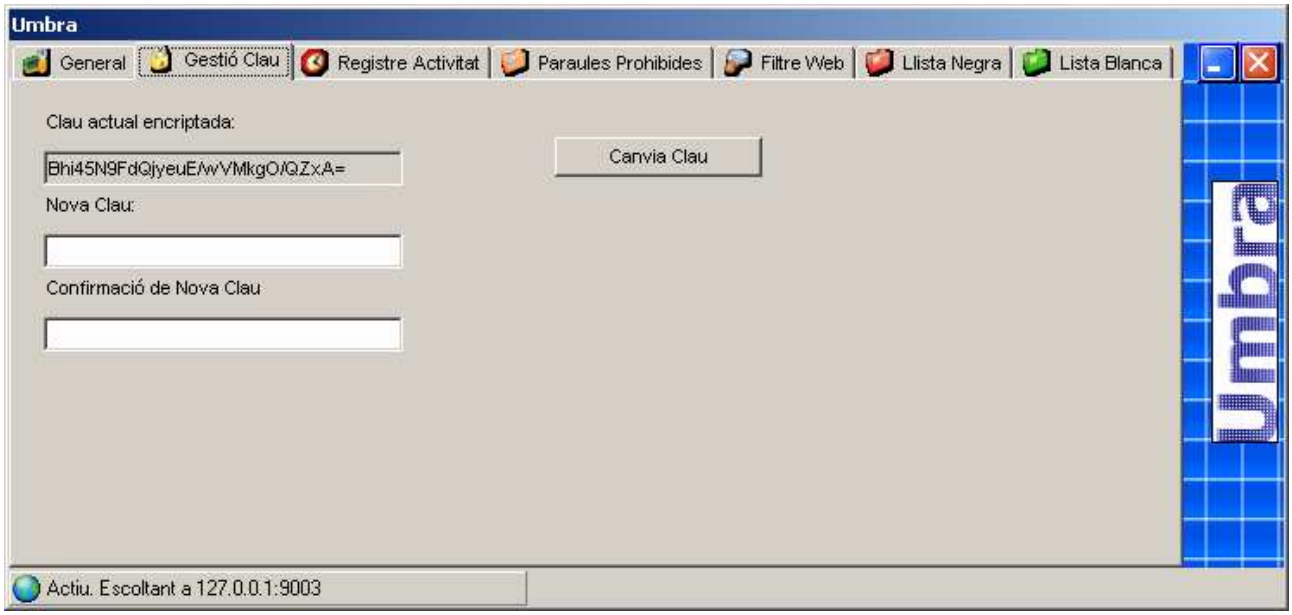
Figura 3-8 Pantalla de configuració del filtratge web



3.9.5 Pantalla de gestió de la clau d'accés

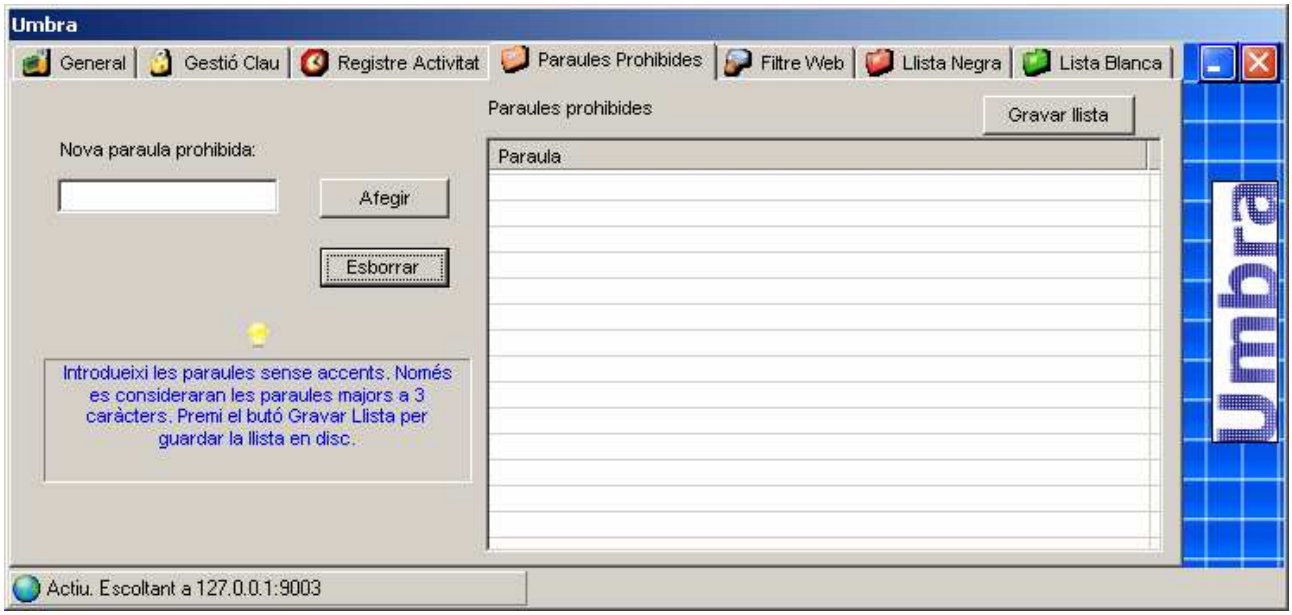
En aquesta pantalla es pot modificar la clau d'accés (la qual permetrà accedir a la configuració). A títol informatiu s'informa de la clau actual encriptada.

Figura 3-9 Pantalla de gestió de la clau d'accés



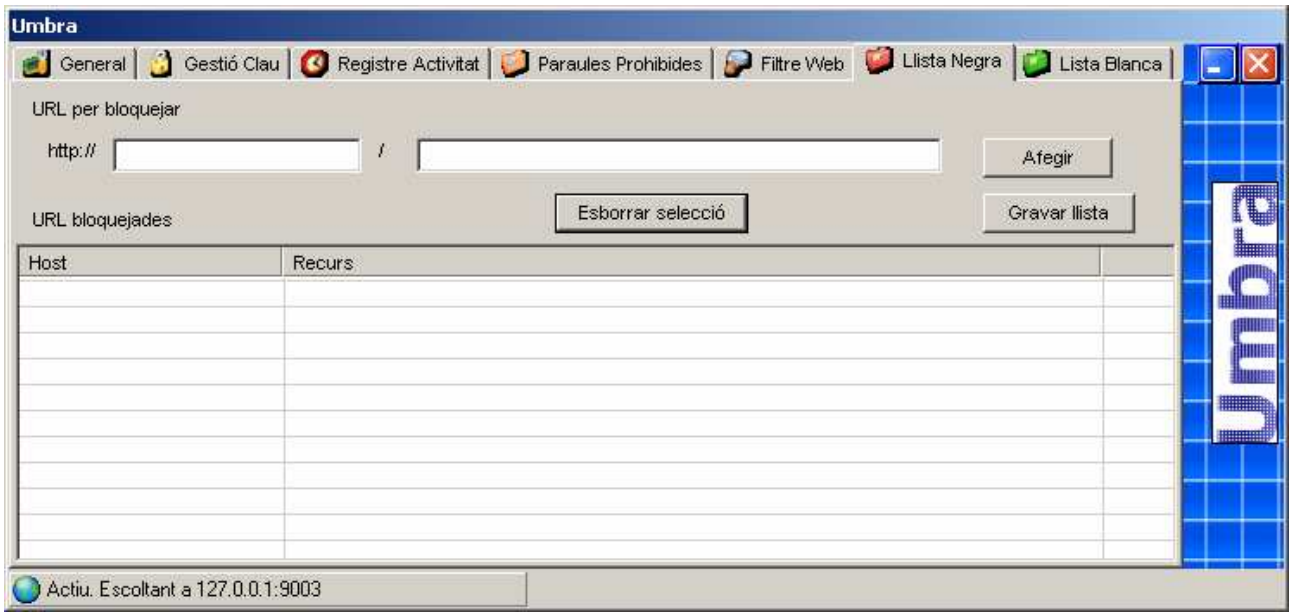
3.9.6 Pantalla de gestió de paraules prohibides

Figura 3-10 Pantalla de gestió de paraules prohibides



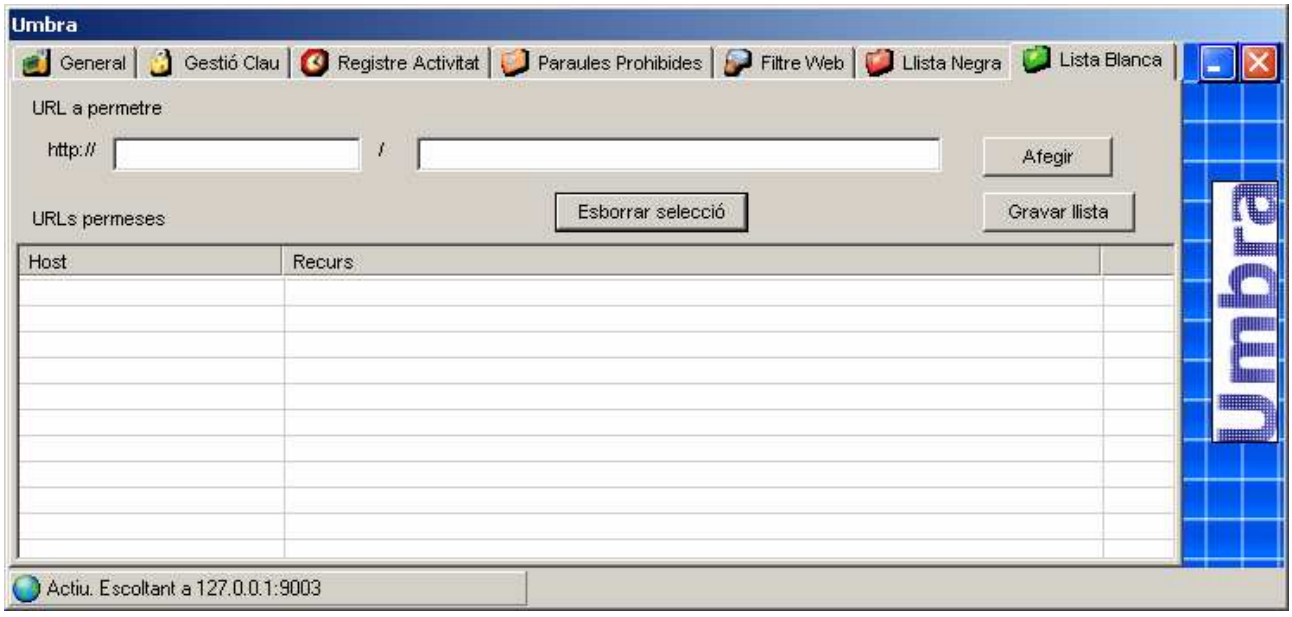
3.9.7 Pantalla de gestió de la llista negra de servidors

Figura 3-11 Pantalla de gestió de la llista negra de servidors



3.9.8 Pantalla de gestió de la llista blanca de servidors

Figura 3-12 Pantalla de gestió de la llista blanca de servidors



4 Disseny orientat a objectes

4.1 Arquitectura

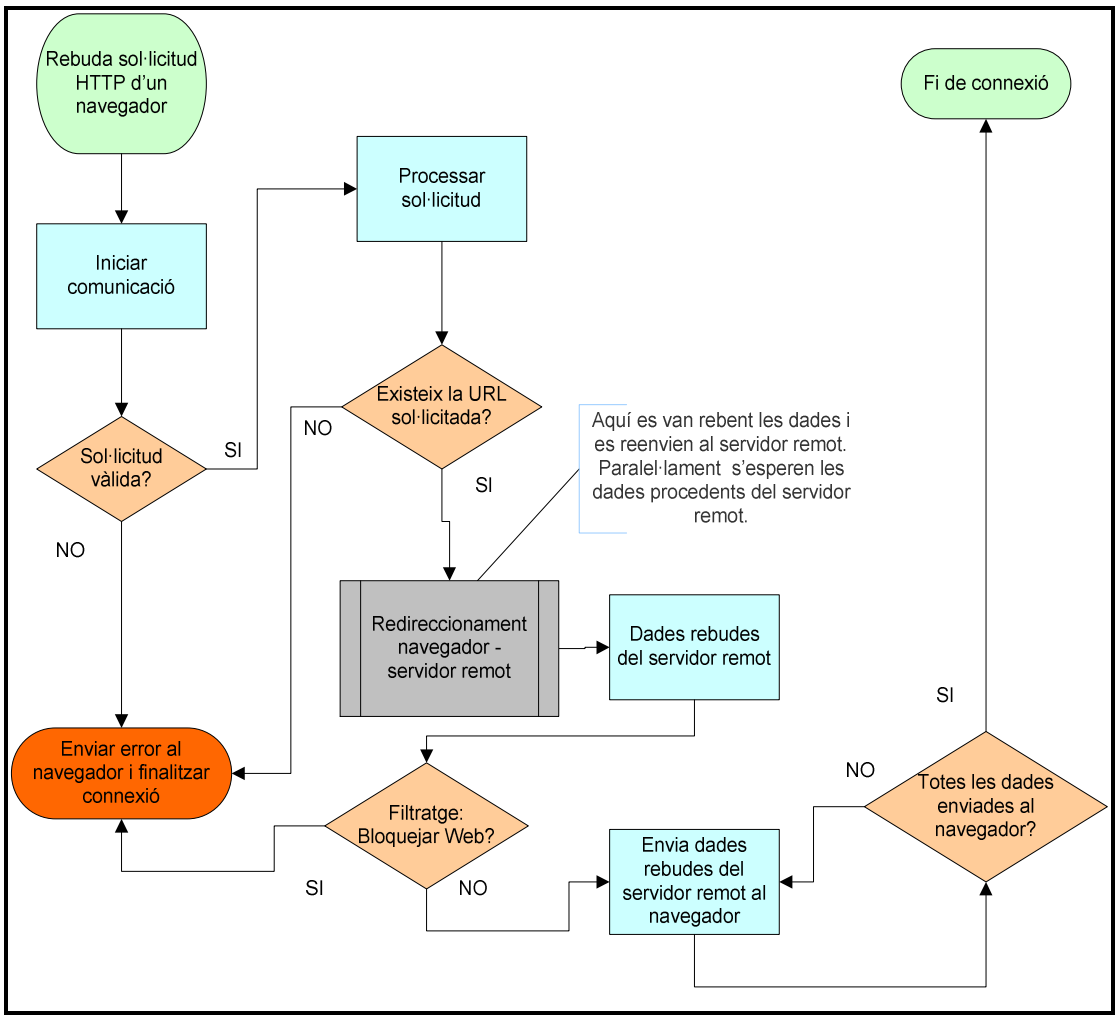
Umbra serà un únic executable configurable via arxius XML i per interfície gràfica i treballarà alhora com un client i un servidor ja que rebrà les peticions HTTP d'un navegador a través d'un port d'escolta, anirà a demanar la pàgina al servidor web remot, realitzarà una tasca de filtratge i, si escau, retornarà una resposta al navegador.

La comunicació entre el navegador i el proxy haurà de ser en la versió 1.0 del protocol HTTP, raó per la qual el navegador haurà de configurar-se convenientment.
A causa de que HTTP 1.0 no admet més que 3 mètodes (GET, POST i HEAD) i les connexions en HTTPS fan servir el mètode CONNECT Umbra serà capaç de tractar aquest tipus de peticions i les deixarà passar sense filtrar els continguts però si filtrarà les URL (sempre que estigui activada la modalidat de filtratge per Llista Negra i Paraules Prohibides).

4.2 Funcionament de l'aplicació

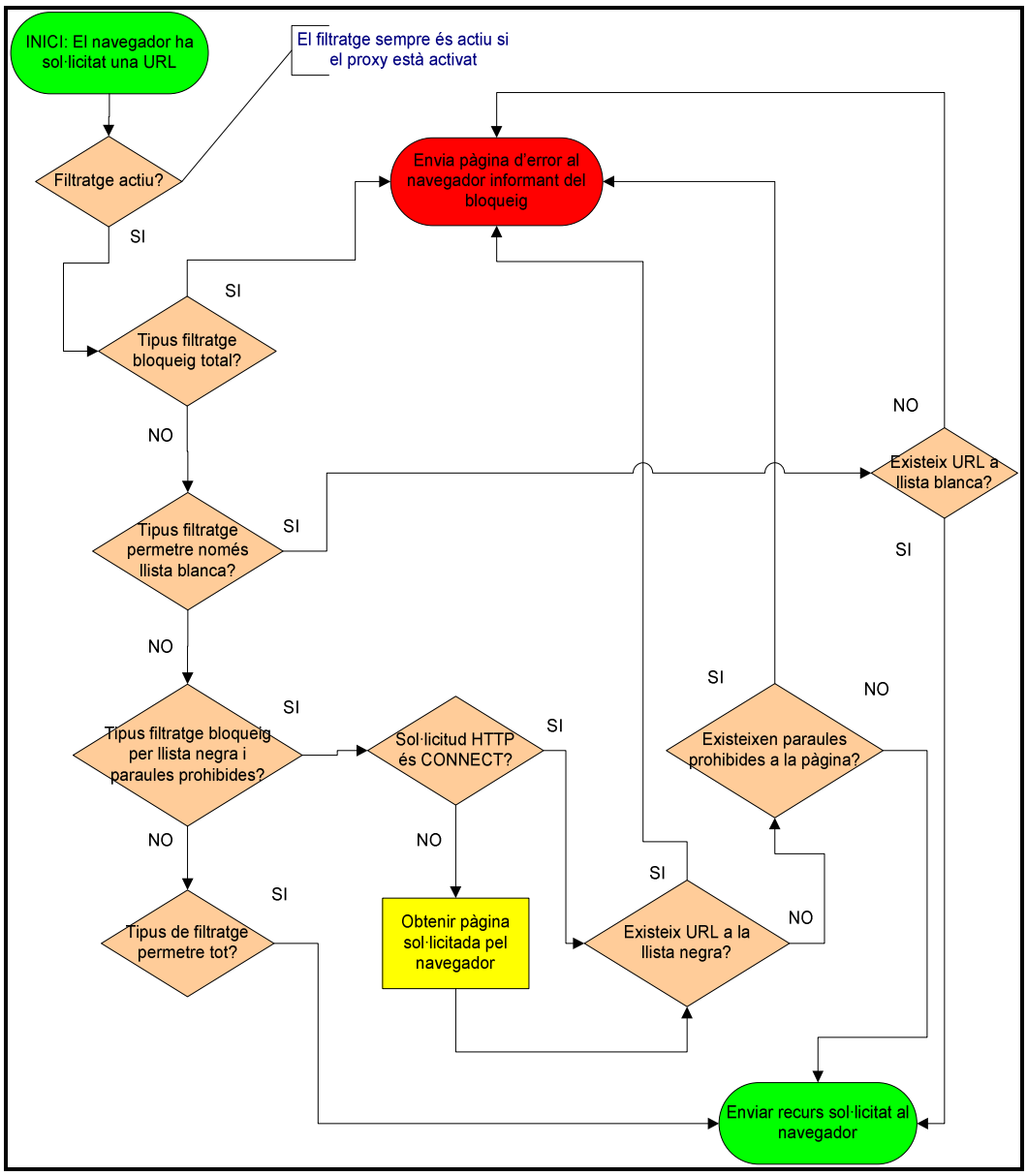
A continuació es presenta el diagrama de fluxe que il·lustra el funcionament general del programa (no s'han tingut en compte les peticions de tipus CONNECT).

Figura 4-1 Diagrama de fluxe del funcionament general



Com es pot veure al diagrama anterior, la lògica de filtratge es troba dins de l'element de decisió "Filtratge: Bloquejar Web?". El diagrama d'aquest mòdul és el següent:

Figura 4-2 Diagrama de fluxe de la lògica de filtratge



4.3 Diagrama estàtic de disseny

El diagrama estàtic de disseny es basa en el realitzat durant l'anàlisi i poseeix més detall que l'anterior.

Els principals canvis que s'han realitzat respecte el diagrama anterior són els següents:

1. Eliminació de la classe ListaPalabrasProhibidas ja que la tecnologia .NET Framework conté una estructura de dades anomenada **DataSet** que és capaç d'enmagatzemar dades obtingudes principalment d'un origen de dades com ara un document XML o una base de dades.
2. Eliminació de la classe ListaServidores pel motiu esmentat anteriorment.

Figura 4-3 Diagrama estàtic de disseny: Cliente i ClienteHTTP

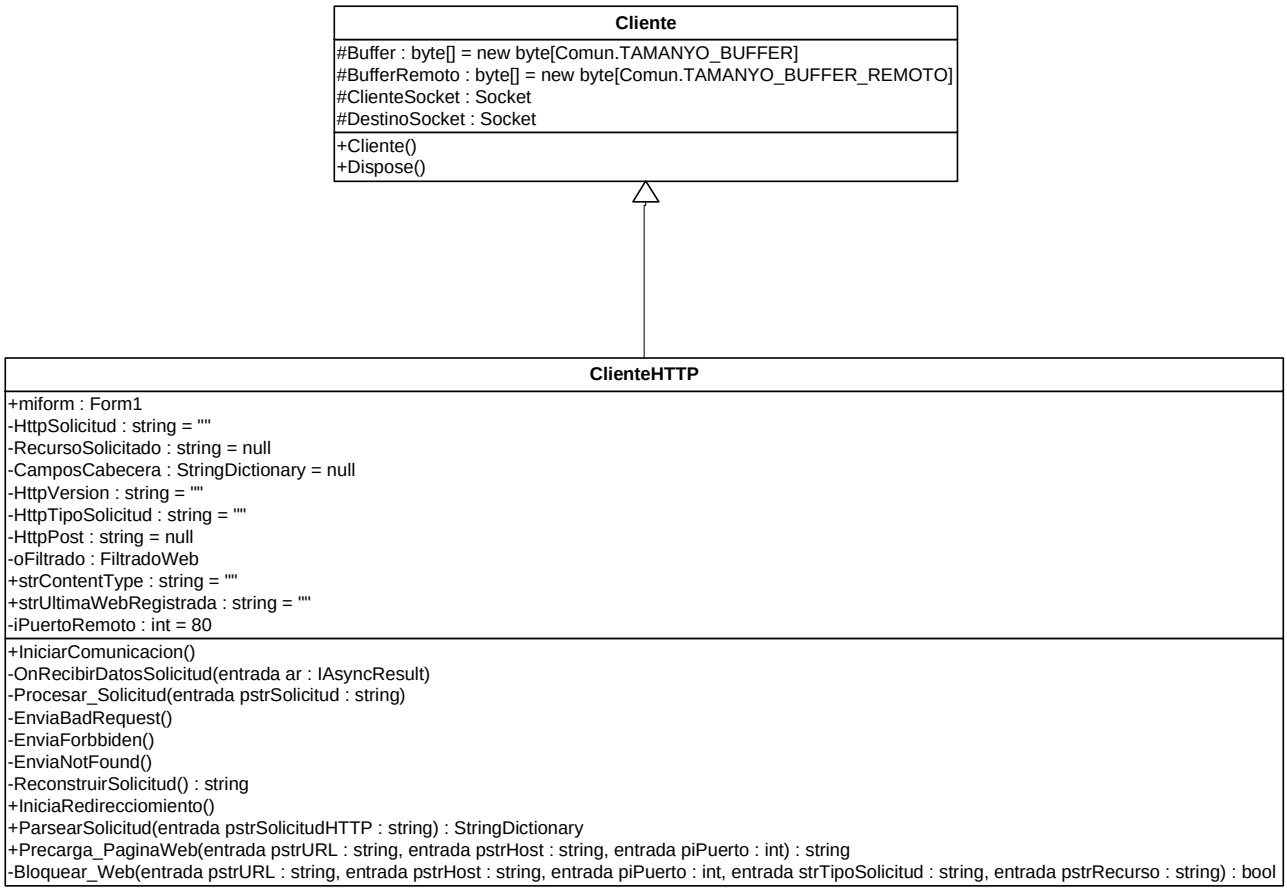
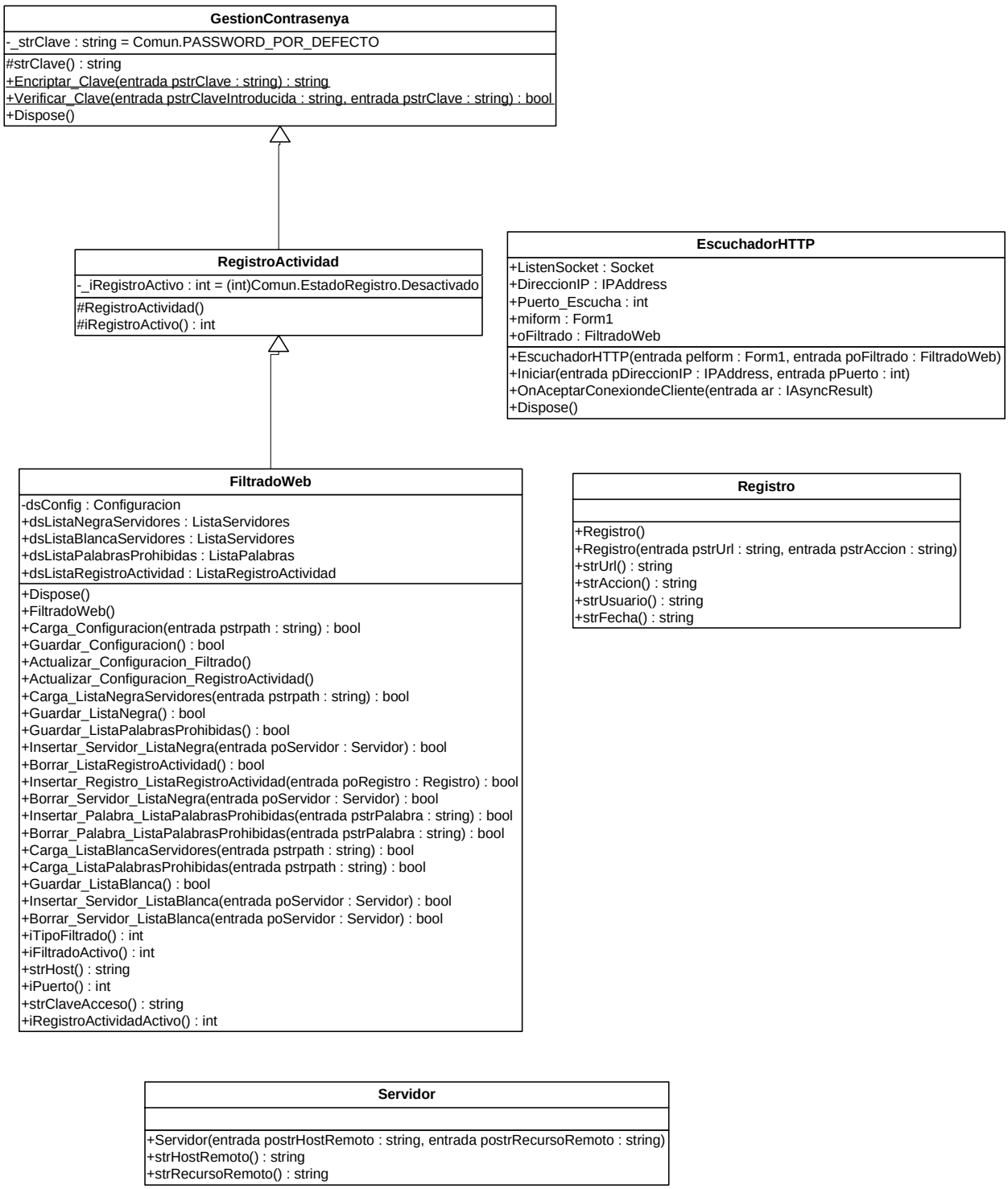


Figura 4-4 Diagrama estàtic de disseny: Comun

Comun
+TAMANYO_BUFFER : int = 4096 +TAMANYO_BUFFER_REMOTO : int = 2048 +NOMBRE_HOST : string = "localhost" +PUERTO_ESCUCHA : int = 9003 +PASSWORD_POR_DEFECTO : string = "Bhi45N9FdQjyeuE/wVMkgO/QZxA=" +NOMBRE_FICHERO_CONFIGURACION : string = "Config.xml" +NOMBRE_FICHERO_LISTA_NEGRA : string = "ListaNegraServidores.xml" +NOMBRE_FICHERO_LISTA_BLANCA : string = "ListaBlancaServidores.xml" +NOMBRE_FICHERO_PALABRAS_PROHIBIDAS : string = "ListaPalabrasProhibidas.xml" +MENSAJE_FILTRO_ACTIVO : string = "Actiu. Escoltant a " +MENSAJE_FILTRO_INACTIVO : string = "Inactiu." +WSAHOST_NOT_FOUND : int = 11001 -HTML_TAG_PATTERN : string = "<.*?>" -HTML_SCRIPT_TAG_PATTERN : string = @"<script>.*?</script>" -HTML_STYLE_TAG_PATTERN : string = @"<style>.*?</style>" -HTML_COMENTARIO_PATTERN : string = @"<!--.*?-->" -HTML_CARACTERES_PATTERN : string = @"&[a-zA-Z]*;" -HTML_META_TAG_PATTERN : string = @"<meta.*?>" +ACCION_WEB_BLOQUEADA : string = "Bloquejada" +ACCION_WEB_PERMITIDA : string = "Permesa" +MAX_NUM_BYTES_RESPUESTA_HTTP : int = 800 +MAX_NUM_REGISTROS_ACTIVIDAD : int = 200
+Parsear_Solicitud_HTTP(entrada pstrHTTP : string, entrada y salida pastrsolicitud : string[]) : bool +EsValida_SolicitudHTTP(entrada pstrSolicitud : string) : bool +Verificar_Primera_Linea_SolicitudHTTP(entrada pstrLinea : string) : bool +Verificar_Cabeceras_SolicitudHTTP(entrada pstrSolicitud : string[]) : bool +AnyadirServidor_ListView(entrada poServidor : Servidor, entrada y salida poListView : ListView) +AnyadirRegistroActividad_ListView(entrada poRegistro : Registro, entrada y salida poListView : ListView) +Existe_Direccion_Web(entrada pdsListaServidores : DataSet, entrada pstrHost : string, entrada pstrRecurso : string) : bool +Elimina_Tags_HTML(entrada pstrPaginaWeb : string) : string +Existe_Palabra_Prohibida(entrada strPaginaWeb : string, entrada pdsPalabras : DataSet) : bool +Existe_Palabra_en_Meta_Tags(entrada pstrPaginaWeb : string, entrada pdsPalabras : DataSet) : bool +Grabar_Fichero_Registro(entrada plvRegistoActividad : ListView, entrada pbMostrarDialogo : bool) : bool

Figura 4-5 Diagrama estàtic de disseny: resta de classes



4.4 Interficie gràfica

4.4.1 Sistema de menús

El programa tindrà un sol menú temporal, en concret un **menú contextual** que només es mostrarà quan estigui minimitzat al *systray*. Si es fa click amb el butó dret del mouse s'obrirà un menú que permetrà accedir a les opcions del programa. Abans de poder accedir es mostrarà la pantalla de control d'accés de forma que no serà possible configurar el programa si no se sap la clau.

Figura 4-6 Menú contextual al *systray*

La resta de menús del programa es troben disponibles en la zona superior de la pantalla (seran permanents), això serà possible gràcies al sistema de *tabs* (pestanyes) el qual permet navegar ràpida i intuitivament per l'aplicació. A la part superior també s'hi podran trobar els butons per tancar el programa i per minimitzar-lo al *systray*.

Figura 4-7 Detall de les opcions de menú permaments

Nota: Per evitar sortir del programa accidentalment s'obrirà un diàleg quan es presioni el botó de sortir.

Figura 4-8 Finestra de confirmació per tancar el programa

4.4.2 Resta de pantalles

La resta de pantalles seran en realitat una de sola implementada a la **classe Form1**.En aquesta classe es trobaran totes les funcions i variables relacionades amb la interfície gràfica mentres que la lògica de filtratge, de xarxa i d'accés a fitxers es troba en les classes del Diagrama estàtic de disseny. Existeix una altra pantalla representada per la classe anomenada **frmControlAcceso** que contindrà una mínima lògica de control d'accés i que serà instanciada des de Form1 en el moment en que s'hagi de sol·licitar la clau d'accés.
La integració de totes les funcionalitats en una mateixa classe (formulari) proporciona un alt grau de senzillesa i no hi ha cap més pantalla que tingui altre tipus de funcionalitats com cerques, modificacions, etc.
Totes les pantalles són les que consten a l'apartat Interfície gràfica del capítol de l'Analisi orientat a objectes.

5 Conclusions

Aquest projecte ha suposat tot un repte personal ja que volia fer servir la tecnologia de Microsoft® .NET Framework amb el llenguatge de programació C# en comptes de Java que està molt més arrelat pel temps que porta en el “mercat” tecnològic.

El fet de que hi hagués molt poca informació a l’abast sobre el tema per l’entorn .NET ha estat tot un estímul en ocasions i una veritable bojeria en d’altres. Tot i així estic satisfet per haver après coses que desconeixia en aquest entorn (tinc certa experiència prèvia) com ara la programació asincrònica i l’actualització dels elements de la interfície gràfica des de threads.

M’hauria agradat afegir-ne més funcionalitats al programa però entre que la informació útil estava molt dispersa i els problemes que he tingut amb coses que desconeixia i que he hagut d’investigar crec que el producte final cumpleix bastant bé les expectatives.

Glossari

Administrador: Usuari que s'identifica amb una contrasenya que li proporciona permissos per configurar totes les opciones del programa.

Contrasenya: paraula secreta que serveix per accedir i configurar les opcions de l'aplicació.

Llista Blanca de Servidors: Llista dels servidors permesos que no han de ser bloquejats.

Llista de Paraules Prohibides: Llista amb totes aquelles paraules l'aparació de les quals en el recurs d'una URL provocarà que es bloquegi l'accés al mateix.

Llista Negra de Servidors: Llista dels servidors prohibits que han de ser bloquejats.

Microsoft® .NET Framework: Model de programació de la plataforma .NET de Microsoft per crear, implementar i executar aplicacions web, aplicacions Windows i serveis web XML. Administra gran part dels detalls d'infraestructura, permetent als desenvolupadors centrar-se a escriure codi de la lògica empresarial pels seus programes. .NET Framework inclou el *Common Language Runtime* o CLR i biblioteques de classes. El CLR és el responsable dels serveis en temps d'execució com la integració de llenguatges, la seguretat i l'administració de memòria, processos i subprocessos. A més té un paper important en temps de desenvolupament ja que té característiques com l'administració de la durada, l'aplicació de noms de tipus segurs, el control d'excepcions entre llenguatges, creació d'enllaços dinàmics, etc.
Per la seva banda les biblioteques de classes proporcionen funcionalitats estàndard com funcions d'entrada i sortida, manipulació de cadenes, administració de seguretat, comunicacions de xarxa, administració de subprocessos i text, característiques de disseny de la interficie d'usuari, accés a bases de dades (classes ADO.NET), manipulació de XML i creació d'aplicacions Windows.
Les biblioteques proporcionen una interfície coherent i comú entre tots els llenguatges compatibles amb .NET Framework (principalment C++, C#, Visual Basic .NET, J#).
Actualment existeixen projecte como MONO (www.mono-project.com) que fa portable .NET a l'entorn Unix.

Proxy HTTP (Proxy web): Programari que permet als clients connectar-se indirectament a servidors web a través d'una xarxa. Els clients es connecten al proxy i fan sol·licituds per accedir a recursos remots. Per la seva banda el proxy els proporciona els recursos connectant-se al servidor remot o bé des de cache.

Registre d'Activitat: registre de l'activitat monitoritzada per l'aplicació a on constarà la URL visitada, usuari, motiu del bloqueig (en cas de que s'hagi bloquejat), data i hora.

Servidor: ordinador remot amb un servidor web que allotja pàgines web les quals es poden localitzar amb una URL cadascuna relativa a aquest ordinador.

URL (*Uniform Resource Locator*): Localitzador Uniforme de Recurs. És una cadena de caràcters, d'acord amb un format estàndard, que s'usa per anomenar recursos com documents i imatges en Internet per la seva localització via WWW (*World Wide Web*).

Usuari: Usuari no administrador del programa sobre el que s'aplica el filtratge de pàgines web.

Bibliografia

Asynchronous Socket Programming in C#: Part I
http://www.codeguru.com/csharp/csharp/cs_network/sockets/article.php/c7695/

Asynchronous Socket Programming in C#: Part II
http://www.codeguru.com/Csharp/Csharp/cs_network/sockets/article.php/c8781/

Client HTTP GET asyncronous
http://samples.gotdotnet.com/quickstart/util/srcview.aspx?path=/quickstart/howto/samples/net/WebRequests/clientGETasync.src

Introduction to socket programming Part 1
http://www.developerfusion.co.uk/show/3918/

Introduction to socket programming Part 2
http://www.developerfusion.co.uk/show/3997/

Windows sockets error codes (Microsoft MSDN Library)
http://msdn.microsoft.com/library/default.asp?url=/library/en-us/winsock/winsock/windows_sockets_error_codes_2.asp

HTTP protocol (World Wide Web Consortium)
http://www.w3.org/Protocols/

HTTP Made Really Easy
http://www.jmarshall.com/easy/http/

Mentalis.org projects
http://www.mentalis.org/soft/projects.qpx

A Tool for Building and Testing Regular Expressions
http://www.codeproject.com/dotnet/expresso.asp

The 30 minute Regex Tutorial
http://www.codeproject.com/dotnet/regextutorial.asp

A Single Instance Application which Minimizes to the System Tray when Closed
http://www.codeproject.com/csharp/singleinstanceapplication.asp

Llamadas Asíncronas
http://www.clikear.com/manuales/csharp/c98.asp

Annex A: Problemes tècnics que han causat retard a la planificació

-Trobar informació, articles, exemples de codi, etc per implementar un proxy fent servir C#.
-Intercepció de les webs abans de que arribin al navegador
-Optimització del proxy a causa del filtratge.
-Debugar codi asincron.
-Trobar un mètode per actualitzar la interfície gràfica d'usuari des d'un subprocés (fil) ja que si es feia com si no es treballés amb fils es produien errors de tipus "subproceso detenido".

Annex B: 2 proves bàsiques del filtre web.

Configuració de prova:

En aquestes condicions es realitzaran un parell de proves per testejar aquesta modalitat.

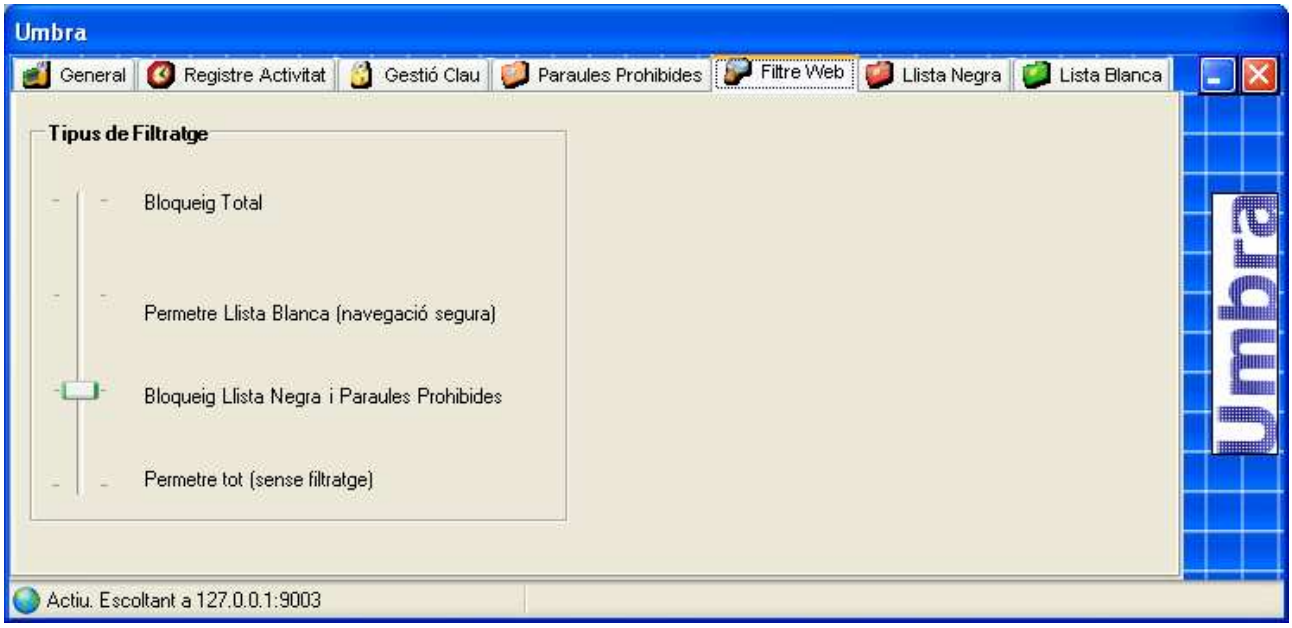


Ilustración 5-1

Prova 1. Filtre per adreça URL:

S'especifica la URL www.terra.es/portada1/portada.html i a continuació es visita www.terra.es

Figura 5-1 Prova de filtratge 1

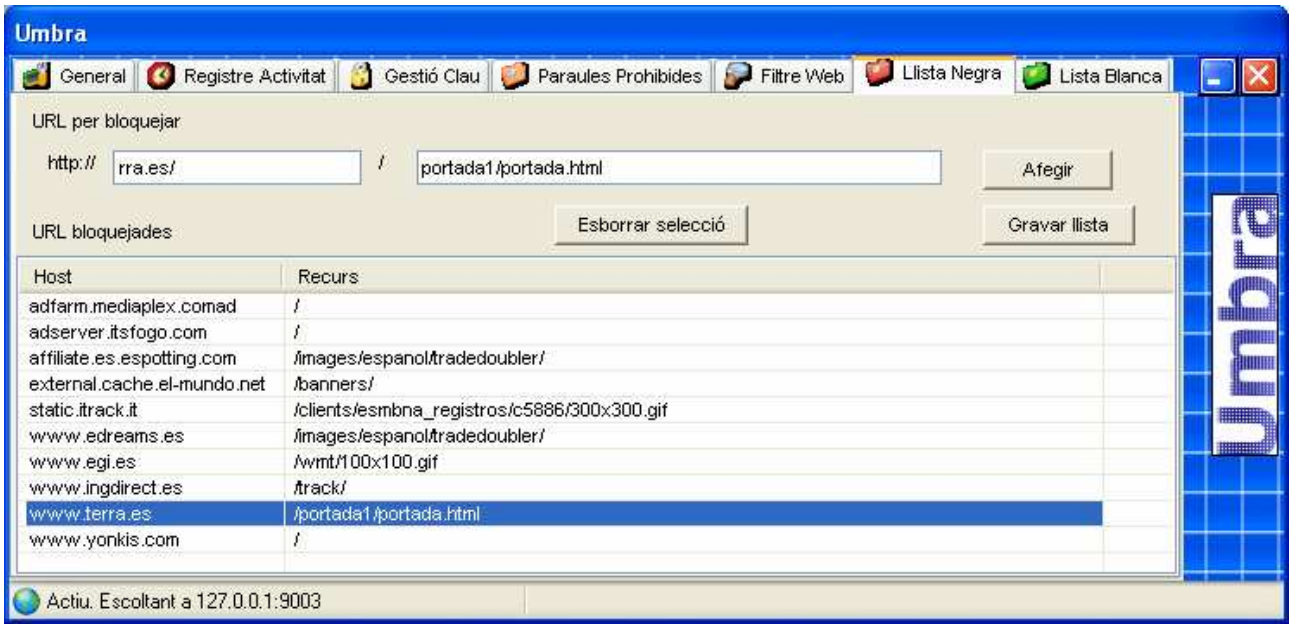
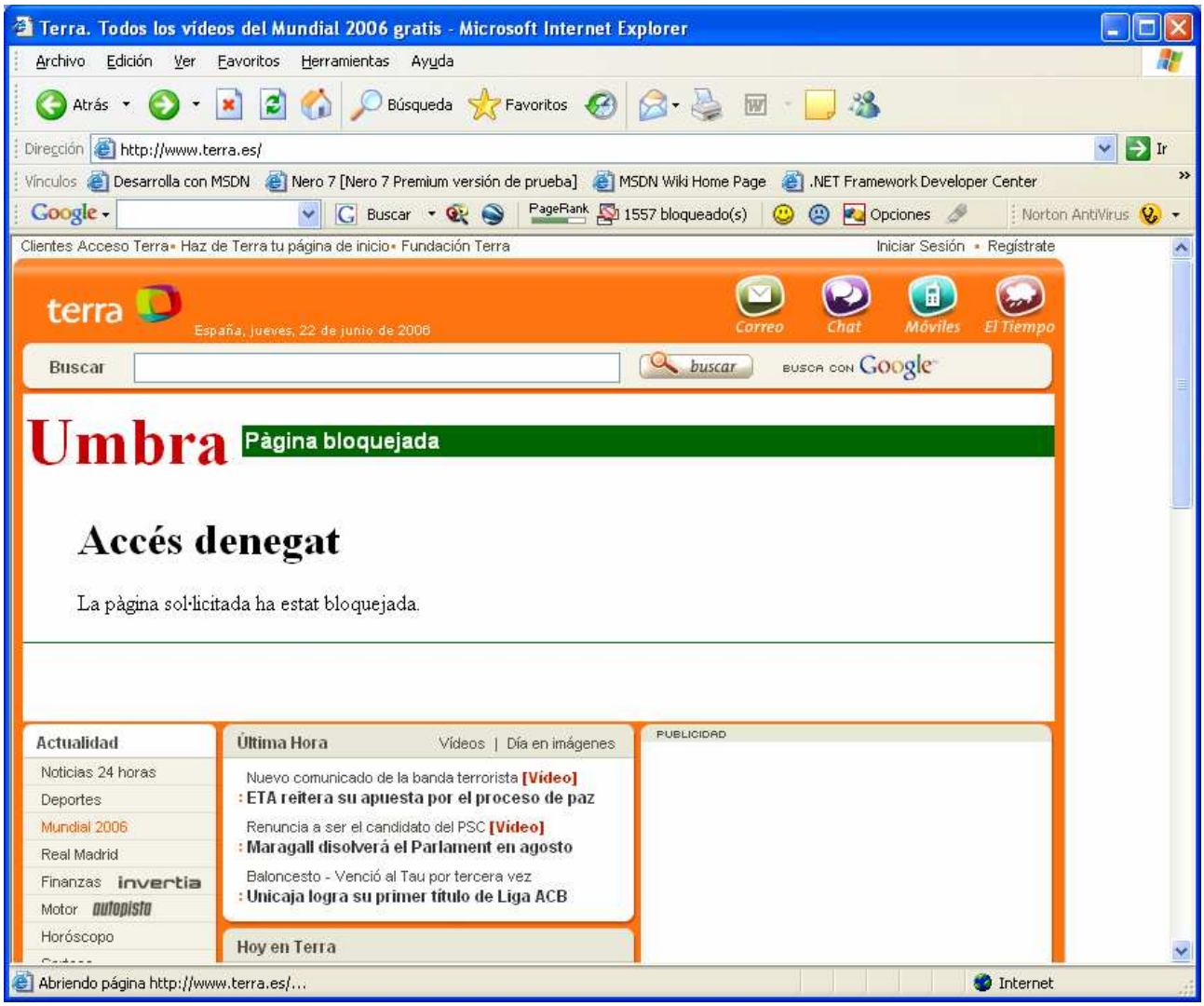


Figura 5-2 Prova de filtratge 1: resultat



Com es pot veure el programa filtra el contigut específic i no tota la web, això permet filtrar banners, imatges, etc, etc.

2. Filtre de contiguts d'una pàgina:

Posem com a paraula prohibida "sex" i visitem www.sex.com

Figura 5-3 Prova de filtratge 2

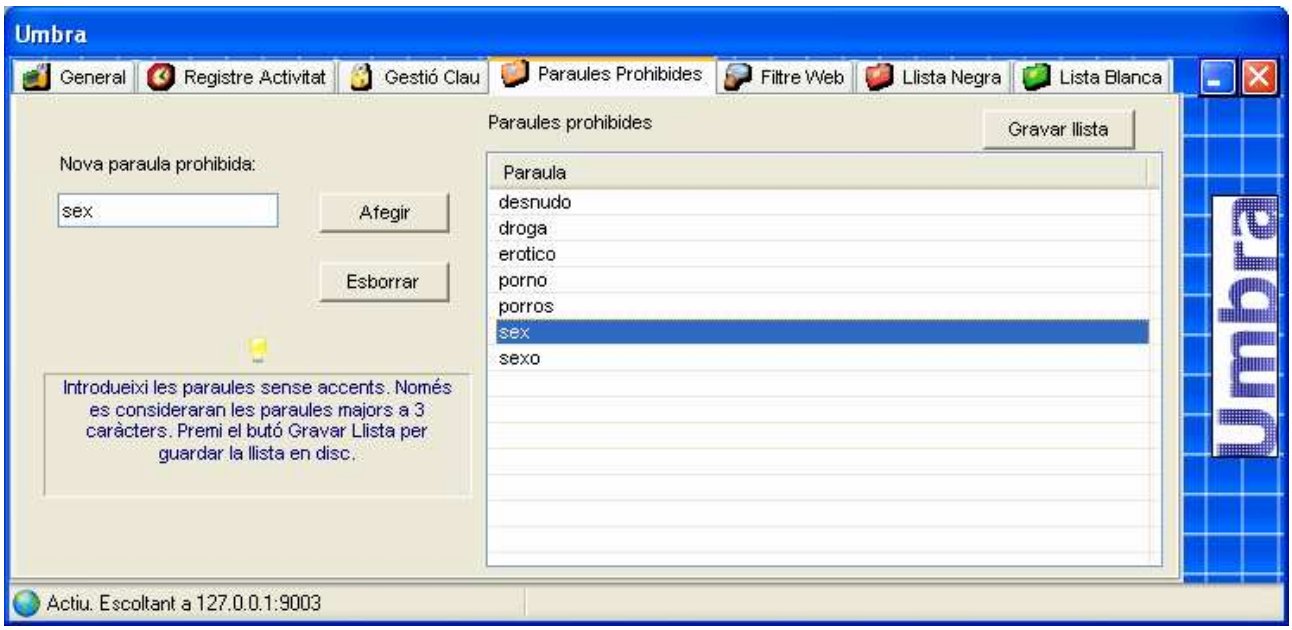
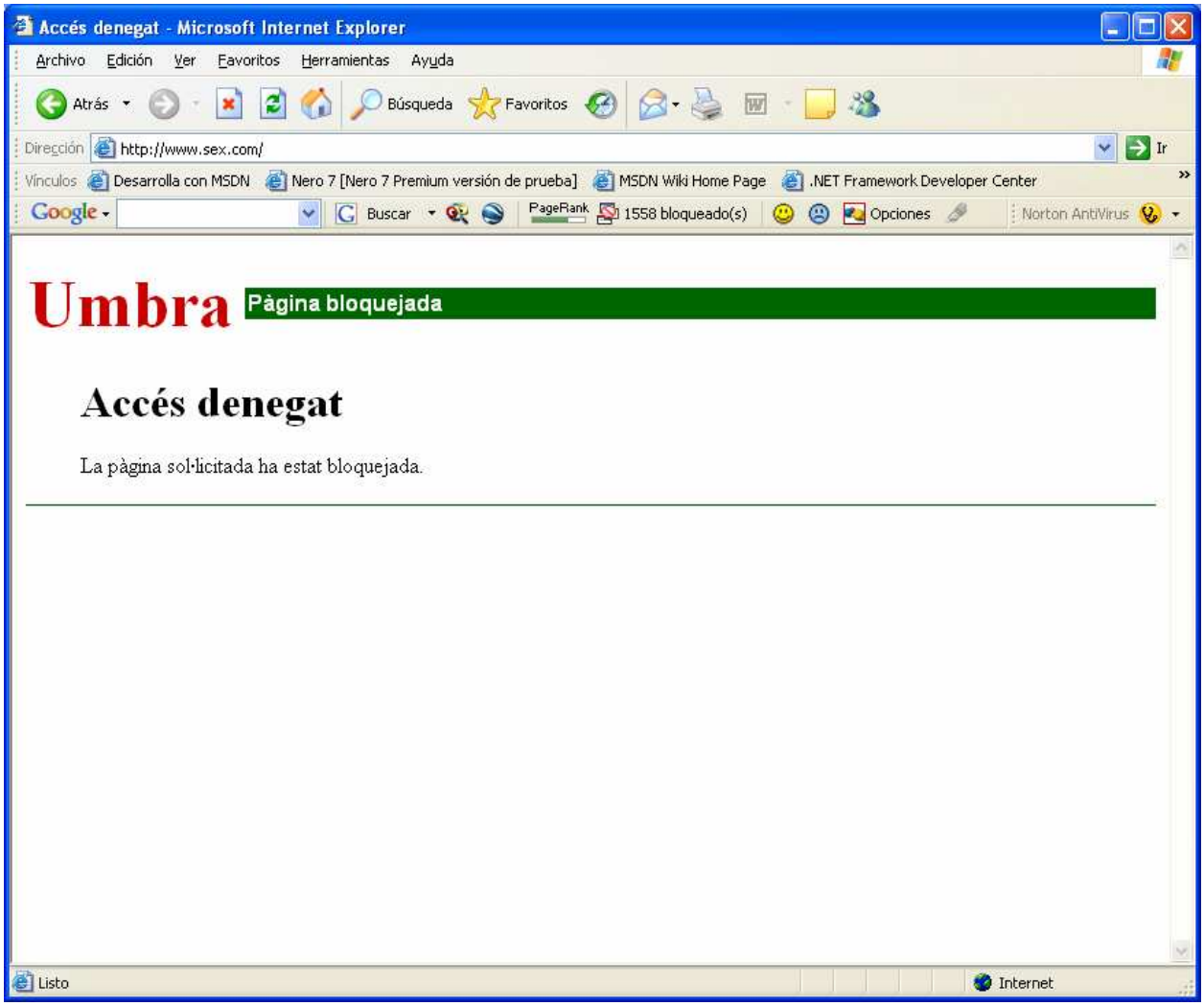


Figura 5-4 Prova de filtratge 2: resultat



La pàgina es bloquejada a causa de que s'ha trobat la paraula sex al contigut de la pàgina.