УДК 621.391

РАСШИРЕННАЯ КЛАССИФИКАЦИОННАЯ МОДЕЛЬ СИСТЕМ ФИЛЬТРАЦИИ КОНТЕНТА В СЕТИ ИНТЕРНЕТ

Каптур В.А., Царёв Р.Ю.

Одесская национальная академия связи им. А.С. Попова,-
65029, Украина, г. Одесса, ул. Ковальская,1.
vadim.kaptur@onat.edu.ua

РОЗШИРЕНА КЛАСИФІКАЦІЙНА МОДЕЛЬ СИСТЕМ ФІЛЬТРАЦІЇ КОНТЕНТУ В МЕРЕЖІ ІНТЕРНЕТ

Каптур В.А., Царьов Р.Ю.

Одеська національна академія зв'язку ім. О.С. Попова,
65209, Україна, м. Одеса, вул. Ковальска, 1.
vadim.kaptur@onat.edu.ua

EXTENDED CLASSIFICATION MODEL OF CONTENT FILTERING SYSTEM IN THE INTERNET

Kaptur V.A., Tsaryov R.Y.

O.S. Popov Odessa national academy of telecommunication,.
1 Kovalska St., 65029, Odessa, Ukraine.
vadim.kaptur@onat.edu.ua

Аннотация. Показана актуальность разработки автоматизированных систем поддержки принятия решений в сфере выбора механизмов фильтрации нецелевого контента сети Интернет. Указано, что ключевой составляющей проекта региональной инициативы Международного союза электросвязи «Создание центра по защите детей в онлайн-среде» по разработке такой автоматизированной системы является база данных существующих технических решений в сфере фильтрации контента. Приведена детализированная классификационная модель систем фильтрации контента в сети Интернет. Модель представлена в виде двух плоскостей: базовых (тип реализации, тип сопровождения, тип совместимости с операционными системами, тип управления, тип внутренней безопасности) и специфических (тип архитектуры, возможность контроля времени работы, объект фильтрации, способ фильтрации, возможность фильтрации шифрованных данных, тип реакции) характеристик. Приведены короткие характеристики каждого из классов систем фильтрации контента, входящих в состав предложенной модели.

Ключевые слова: фильтрации контента, база данных существующих решений, классификационная модель, защита детей в онлайн-среде.

Анотація. Показано актуальність розробки автоматизованих систем підтримки прийняття рішень у сфері вибору механізмів фільтрації нецільового контенту мережі Інтернет. Зазначено, що ключовою складовою проекту регіональної ініціативи Міжнародного союзу електрозв'язку «Створення центру із захисту дітей в онлайн-середовищі» з розробки такої автоматизованої системи є база даних існуючих технічних рішень у сфері фільтрації контенту. Наведена деталізована класифікаційна модель систем фільтрації контенту в мережі Інтернет. Модель показана у вигляді двох площин: базових (тип реалізації, тип супроводу, тип сумісності з операційними системами, тип управління, тип внутрішньої безпеки) і специфічних (тип архітектури, можливість контролю часу роботи, об'єкт фільтрації, спосіб фільтрації, можливість фільтрації шифрованих даних, тип реакції) характеристик. Наведено короткі характеристики кожного з класів систем фільтрації контенту, що входять до складу запропонованої моделі.

Ключові слова: фільтрація контента, база даних існуючих рішень, класифікаційна модель, захист дітей в онлайн-середовищі.

Abstract. The urgency of the development of automated systems to support decision-making in the choice of filtering mechanisms inappropriate content. It pointed out that a key component of the project of a regional initiative of the International Telecommunication Union "Creation of the center for the child online protection" for the development of this automated system is a database of existing solutions in the area of content filtering. Shows detailed classification model of content filtering systems in the Internet. The model is presented in the form of two planes: the base (the type of implementation, the type of support, type of compatibility with operating systems, the type of control, type of internal security) and specific (type of architecture, the ability to control operating time, the object of filtration method for filtering, the ability to filter encrypted data, type of reaction) characteristics. Presents short characteristics of each class of the content filtering systems that are part of the proposed model.

Key words: content filtering, database of existing solutions, classification model, child online protection.

Современные информационные технологии предоставляют уникальные возможности для доступа к практически неограниченному объему разнотипной информации. Человек, использующий сеть Интернет, становится полноправным участником глобального информационного социума, в котором его подстерегает немало опасностей. Интернет, как и любой другой инструмент, созданный человечеством, может нести как пользу, так и вред.

Некоторая часть информации (порнография, сцены насилия, пропаганда наркотиков, алкоголя, терроризма, нацизма и т.п.), циркулирующая в сети Интернет, может нанести психологическую травму, привить неправильные морально-этические качества и даже сделать человека жертвой психологического запугивания, притеснения и сексуального преследования. Для защиты человека от негативной стороны использования сети Интернет рекомендуется использовать комплексный подход, который включает организационно-педагогические и технические меры [1].

Основной технической мерой защиты человека от негативной информации в сети Интернет служит техническая фильтрация информации. Многообразие характеристик систем технической фильтрации информации порождает множество вариантов их построения и, как следствие, огромное количество программных и программно-аппаратных решений, предназначенных для блокирования доступа к информационным ресурсам на различных уровнях и в сетях разного типа.

Появление значительного количества программных и программно-аппаратных решений, предназначенных для блокирования доступа к информационным ресурсам в компьютерных сетях различного типа поставило перед пользователями новую задачу – задачу выбора наиболее подходящего с технической и экономической точек зрения решения для конкретной ситуации (дома, школы, офиса и т.д.) [2].

Выбор наиболее подходящего для конкретной ситуации решения, как правило, основывается на комплексном анализе множества факторов, таких как: область применения (персональные компьютеры дома, несколько компьютеров в школе, большая компьютерная сеть в учебном заведении, смартфоны и планшеты, подключенные к сети оператора и т.д.), способ подключения к сети Интернет (один или несколько каналов доступа), наличие специалистов для установки и поддержки решения, наличие «свободной» техники для установки нового решения, желаемая политика доступа к сети Интернет и так далее.

Очевидно, что проведение подобного анализа требует привлечения высококвалифицированных специалистов в области фильтрации контента, что не всегда является возможным, особенно в реалиях общеобразовательных учебных заведений региона СНГ. Учитывая это на Всемирной конференции по развитию электросвязи 2014 года (Дубай, ОАЭ) была принята региональная инициатива «Создание центра по защите детей в онлайн-среде региона СНГ» [3]. Одним из базовых проектов этой инициативы стал проект по созданию единой базы данных с информацией про существующие технические решения и системы выбора оптимальной (для конкретной организации) системы фильтрации контента. Одним из первых шагов на пути реализации данного проекта является разработка классификационной модели систем фильтрации контента, использующихся сегодня в сети

Интернет. На предыдущих этапах исследования была предложена обобщённая модель фильтрации web-контента в сети Интернет, которая позволяет провести классификацию существующих средств, видов, методов и подходов к фильтрации [4]. Однако, данная модель не отражает того множества характеристик систем фильтрации контента, которые могут послужить критерием выбора наиболее подходящей системы со стороны пользователя.

Целью данной работы является разработка детализированной классификационной модели систем фильтрации контента в сети Интернет, как основы базы данных существующих технических решений.

Множество систем фильтрации контента, созданных разработчиками программного обеспечения и производителями телекоммуникационного оборудования по всему миру, может быть описано на основании классификационной модели. Предлагаемая классификационная модель (рис. 1) представлена в виде двух плоскостей – плоскости базовых характеристик и плоскости специфических характеристик. В плоскости базовых характеристик для классификации систем фильтрации контента использует следующие критерии: тип реализации; тип сопровождения; тип совместимости с операционными системами; тип управления; тип внутренней безопасности.

В плоскости специфических характеристик используют следующие критерии: тип архитектуры; возможность контроля времени работы; объект фильтрации; способ фильтрации; возможность фильтрации шифрованных данных; тип реакции (способ работы).

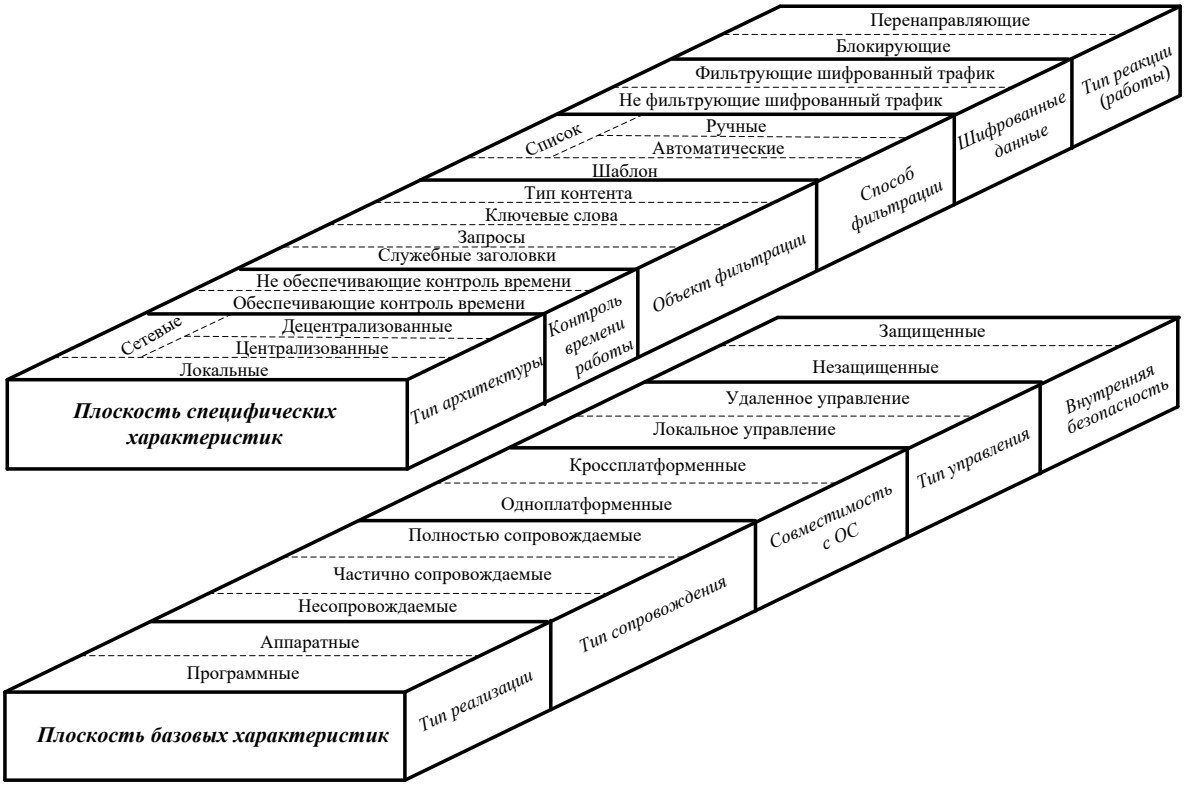

Рисунок 1 – Классификационная модель систем фильтрации контента

Тип реализации (уровень базовых характеристик):

– Программные. Системы данного класса реализованы в виде программного продукта, который устанавливается либо непосредственно на устройство пользователя либо на специально выделенный сервер. Достоинством систем данного класса является их относительная простота установки, настройки и эксплуатации;

– Аппаратные. Системы данного класса реализованы в виде аппаратного продукта, который устанавливается как отдельное сетевое устройство. Недостатком систем данного класса является высокая стоимость решения, трудоемкость и сложность настройки.

По типу сопровождения (уровень базовых характеристик):

– Полностью сопровождаемые системы – системы данного класса централизовано и регулярно обновляются с целью улучшения эффективности их работы и удобства эксплуатации. Так же для данных систем регулярно обновляются базы стоп-листов (черных/белых списков) ресурсов, вирусов;

– Частично сопровождаемые системы – для систем данного класса централизовано и регулярно обновляются только базы стоп-листов (черных/белых списков) ресурсов, вирусов. Обновление функциональных возможностей происходит в случае выявления критических ошибок в работе системы;

– Несопровождаемые системы (US) – централизованное обновление систем данного класса не проводится. Разработчик системы, как правило, не осуществляет обновлений и пользователь должен сам решать эти задачи.

По совместимости с операционными системами (уровень базовых характеристик):

– Одноплатформенные. Системы данного класса предназначены для работы с устройствами, которые находятся под управлением операционной системы определенного типа, например, только Windows или только Unix;

– Кроссплатформенные. Системы данного класса поддерживают работу с несколькими ОС, например, система может работать на устройствах под управлением ОС Windows и на устройствах под управлением ОС Android.

По типу управления (уровень базовых характеристик):

– Системы локального управления – управление системами данного класса осуществляется «локально» изнутри защищаемого объекта;

– Системы удаленного управления – управление системами данного класса возможно не только «локально» изнутри защищаемого объекта.

По типу внутренней безопасности (уровень базовых характеристик):

– Защищенные системы – у систем данного класса настройки системы (параметры фильтрации) защищены паролем, не каждый пользователь, который имеет доступ к системе, может изменять режим ее работы;

– Незащищенные системы – у системы данного класса настройки системы (параметры фильтрации) не защищены паролем, любой пользователь имеющий доступ к системе может изменять режим ее работы.

По типу архитектуры (плоскость специфических характеристик):

– Локальные – предназначены для организации фильтрации контента только на одном устройстве (компьютер, смартфон, планшет);

– Сетевые – предназначены для организации фильтрации контента для устройств, объединенных в сеть:

– Централизованные – предполагает наличие единой, для всех устройств, политики фильтрации, которая осуществляется центральным элементом (устройством). Достоинством является простота организации и относительно низкая стоимость, недостаток – при выходе из строя центрального элемента система становится полностью не работоспособной;

– Децентрализованные – фильтрация осуществляется группой элементов (устройств) распределенных по сети, каждый из которых может реализовывать свою собственную политику фильтрации, центральный элемент осуществляет функцию согласования и координации работы остальных элементов. Достоинством является гибкость настройки

(можно организовать уникальные политики фильтрации для отдельных пользователей, группы пользователей) и высокая надежность (отказ одного из элементов не приводит к выходу из строя всей системы). Недостатки – более высокая стоимость решения, требует больше трудозатрат на настройку и последующую эксплуатацию.

По возможности контроля времени работы (плоскость специфических характеристик):

– Системы, обеспечивающие контроль времени работы в сети Интернет – позволяют контролировать время, проведенное в сети Интернет пользователями;

– Системы, не обеспечивающие контроль времени работы в сети Интернет.

По объекту фильтрации (плоскость специфических характеристик):

– Фильтрация по служебным заголовкам (полям) – системы данного класса позволяют осуществлять фильтрацию на базе таких служебных данных, как тип протокола, номер порта и т.д.;

– Фильтрация по запросам к поисковым системам – анализируется содержимое запросов пользователей отправляемых к поисковым системам на соответствие заданным регулярным выражениям. Проверяется как запрос пользователя, так и ответ ресурса;

– Фильтрация по ключевым словам (морфологический анализ) – анализируется содержимое ресурса (сайта) на наличие определенных слов или словосочетаний. Если контент содержит указанные ключевые слова/словосочетания, то ресурс блокируется;

– Фильтрация по типу контента – блокирует доступ/загрузка контента заданного типа – видеоконтента, аудиоконтента, изображений, документов, архивов и исполняемых файлов и т.д.

По способу фильтрации (плоскость специфических характеристик):

– Фильтрация по спискам – решение о доступе к ресурсу принимается на основе анализа черных (запрещенных)/белых (разрешенных) списков. В зависимости от способа формирования списков можно выделить два подкласса: системы с ручным формированием списков; системы с автоматическим формированием списков.

– Фильтрация по шаблонам – решение о доступе к ресурсу принимается на основе анализа шаблона, если анализируемый ресурс совпадает с заданным шаблоном, то он блокируется/разрешается.

По возможности фильтрации шифрованных данных (плоскость специфических характеристик):

– Системы, позволяющие фильтровать шифрованные данные – системы данного класса предоставляют возможность фильтровать данные, которые передаются по шифрованным каналам (протоколы HTTPS, SSH);

– Системы, не позволяющие фильтровать шифрованные данные.

По типу реакции (способу работы) (плоскость специфических характеристик):

– Блокирующие – системы данного класса при обнаружении попытки доступа к запрещенному контенту/ресурсу блокируют доступ к этому контенту/ресурсу;

– Перенаправляющие – системы данного класса при обнаружении попытки доступа к запрещенному контенту/ресурсу переадресовывает пользователя на специальную страницу, где объясняется, почему не может быть предоставлен доступ к запрошенному контенту/ресурсу или где может быть размещена любая другая информация, заданная администратором системы.

Выводы и результаты:

1. Предложенная классификационная модель отражает как базовые (характеризующие любую автоматизированную систему), так и специфические для систем фильтрации контента характеристики.

2. Модель может быть положена в основу базы данных с информацией про существующие технические решения системы фильтрации контента как составной части проекта по созданию автоматизированной системы поддержки принятия решений в сфере выбора механизмов фильтрации нецелевого контента сети Интернет.

3. Дальнейшие исследования должны быть направлены на разработку методики выбора оптимальной (для конкретной организации) системы фильтрации контента.

ЛИТЕРАТУРА:

1. Каптур В. Сучасний стан та перспективи розвитку методів фільтрації контенту в телекомунікаційних мережах / В. Каптур // Безпека інформації. – 2014. – Т. 20. – № 2. – С. 113-119. –Режим доступу: http://nbuv.gov.ua/j-pdf/bezin_2014_20_2_3.pdf.
2. Каптур В.А. Методика визначення найбільш ефективного способу організації системи фільтрації нецільового контенту в мережі організації / В.А. Каптур, К.Д. Гуляєв, П.С. Кравченко // Наукові праці ОНАЗ ім. О.С. Попова. – 2012. – № 1. – С. 47 – 52.
3. Всемирная конференция по развитию электросвязи 2014 года: Финальный отчёт [Электронный ресурс]. – Режим доступа: http://www.itu.int/pub/D-TDC-WTDC-2014.
4. Каптур В.А. Узагальнена класифікаційна модель фільтрації контенту в мережі Інтернет / В.А. Каптур // Збірник наукових праць Військового інституту телекомунікацій та інформатизації НТУУ "КПІ". – 2011. – № 1. – С. 65 – 70.

REFERENCES:

1. Kaptur, V. "Current Status and Prospects of the Content Filtering Methods in the Telecommunication Networks." Ukrainian Scientific Journal of Information Security 20.2 (2014): 113-19. Web. <http://nbuv.gov.ua/j-pdf/bezin_2014_20_2_3.pdf>.
2. Kaptur, V.A., K.D. Guliaev, and P.S. Kravchenko. "Selection of the Content Filtering Systems Configuration for Corporate Networks." Proceedings of the O.S. Popov ONAT 1 (2012): 47-52. Web.
3. "World Telecommunication Development Conference 2014: Final Report."ITU. N.p., n.d. Web. http://www.itu.int/pub/D-TDC-WTDC-2014.
4. Kaptur, V. "General classification model of content filtration in the Internet." Proceedings of the VITI "KPI" 1 (2011): 65-70.