**ТЕМА 6. ИЗВЛИЧАНЕ НА ЗНАНИЯ ОТ УЕБ ИЗТОЧНИЦИ   
(WEB MINING - WM)**

**Съдържание**

6.1. Същност на процеса Web Mining.

6.2. Етапи на процеса на Web Mining

# 6.3. Извличане на знания от съдържанието на уеб източници

6.4. Извличане на знания от структурата на уеб източници

6.5.Извличане на знания от информацията за ползваемостта на уеб източници

# 6. 1. Същност на процеса Web Mining

Интернет се обособи като най-богатия източник на лесно достъпна информация. В мрежата има голямо количество документи, данни аудио и видео файлове, от които чрез подходяща обработка биха могли да се извлекат нови и полезни за бизнеса знания. Познанието се генерира не само от съдържанието на самите страници, но също и от уникалните характеристики, структурата на уеб сайтовете и от информацията за достъпа до тях.

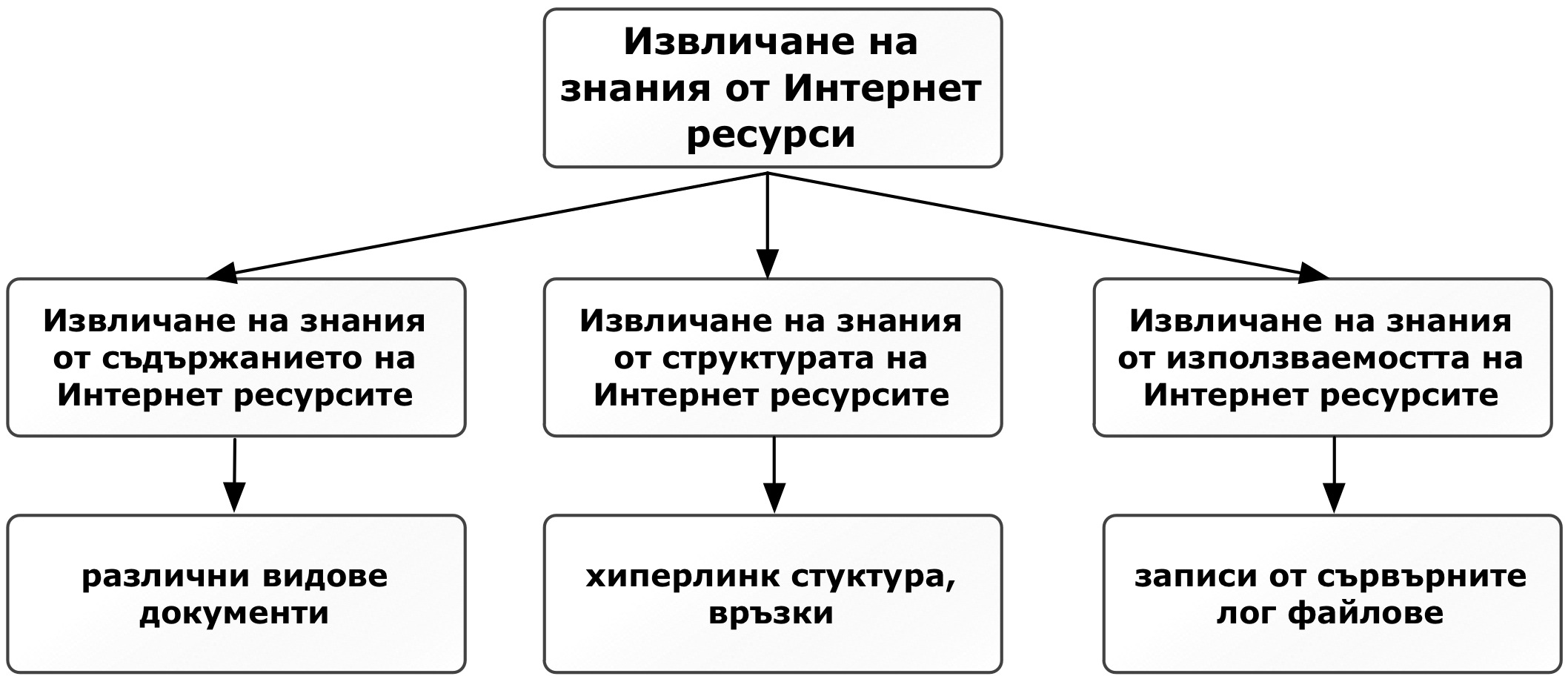
Основните характеристики, които определят Интернет като подходящ източник за нови знания са:

* големият обем и обхват информация. Световната мрежа Интернет, с бурните си темпове на развитие обхваща все повече потребители, приблизително 2,1 млрд. души.
* Интернет служи за пренасяне на голямо разнообразие от информационни ресурси, които съдържат и голям обем полезна за икономическите единици информация.
* свързаността на информацията, чрез използването на хипервръзки. Интернет служи за пренасяне на голямо разнообразие от информационни ресурси и услуги, най-известни сред които са взаимосвързаните хипертекстови документи. Известно е, че хипервръзките препращат към други документи, които имат връзка с разглеждания, освен това те могат да се проследяват автоматично от специални програми за търсене и индексация, наречени уеб spider или уеб crawler.
* хетерогенният характер на информацията. В Интернет основните източници на информация са уеб сайтовете, които включват в своето съдържание информация под формата на текст, таблици, изображения, мултимедия (аудио, видео).
* динамичността на информацията. Публикуваната информация в Интернет се обновява непрекъснато, появяват се актуални гледни точки за проблемите, което в много случаи се явява източник на нови знания.

Натрупването на все повече информация в уеб пространството е предпоставка за прилагането на познатите техники за извличане на знания от интернет източници. Извличането на знания от уеб ресурси се нарича Web Мining (WM). В литературата този термин се използва за първи път от Ециони през 1996 г. Tой определя WM като процес на използване на DM техниките за изследване и извличане на информация от интернет документи и услуги.По-късно понятието WM се обогатява, като към него се включват и техниките за изследване и анализ на данните за използваемостта на уеб ресурсите. Други автори дефинират процеса на Web Мining, като “разкриване на закономерности, зависимости, шаблони в съдържанието, структурата и начина на използването на уеб документите, които показват как да се приложат DM методите и моделите върху уеб базирани данни”.

Като обобщение на цитираните определения може да се счита най-популярното в литературата становище за WM - “откриване и анализ на полезна информация от Световната мрежа”.

В зависимост от основните типове интернет ресурси, които се използват в процеса на WM, в литературата са дефинирани три основни вида извличане на знания (фиг. 6.1):



Фиг. 6.1. Видове WM

* **Web content mining (WCM).** Определя се като процес на извличане на полезни знания от съдържанието на уеб документите. Например, чрез WCM може: автоматично да се класифицират и групират (клъстеризират) уеб страници, според техните теми; да се изследва какво търсят потребителите в Интернет и да се откриват характерни шаблони за търсене на зададената група от хора; да се откриват модели в уеб страниците на базата на описанията на продуктите, съобщения с мнения от форуми и др. Освен това WCM се използва за подобряване намирането и филтрирането на информация, категоризиране на документи, намиране на подобни страници върху различни сървъри, идентифициране на теми, представени в различни уеб документи.
* **Web structured mining (WSM).** Процес на откриване на полезни знания в хипервръзките, чрез изследване на структурата от входящи и изходящи връзки (топологията) в уеб. Прилага се теорията на графите за анализиране на даден сайт и структурата му от връзки. Изследванията се групират в две направления: извличане на модели от хипервръзките на уеб и изследване на структурата на документа. Най-често се използва дървовидна структура за анализ и описание на таговете (HTМL, XML) в уеб страницата. Например от връзките може да се открият важни уеб страници, общности от потребители, които споделят общи интереси и др.
* **Web usage mining (WUM).** Процес на откриване на потребителски модели на базата на данните за използване на интернет ресурси. Осигурява намиране на шаблони в потока от кликвания и асоциирани данни, събрани или генерирани като резултат от потребителската интерактивност с уеб сайтовете. За целите на WUM обикновено се анализират записите за логванията към сървърите и се получава информация за наборите от страници, обекти или ресурси, които често се достъпват от групи потребители с общи интереси. Получените знания са полезни за процеса на персонализация, за търсещите машини, за усъвършенстване на системите и др.

# 6.2. Етапи на процеса на Web Mining

Въпреки, че извличането на данни от уеб източници води началото си от DM, трябва да се отбележи, че двата процеса не са идентични. Основната причина за това е, че традиционно извличането на знания използва структурирани данни, съхранявани в складове от данни, релационни таблици, електронни таблици или текстови данни, но в таблична форма. Уеб данните трябва първо да бъдат открити и след това да се сведат до вид подходящ за анализ, тъй като те по своето същество са полу- или неструктурирани.

В литературата процесът на Web Mining също може се срещне декомпозиран по различен начин в зависимост от това дали подзадачите са дадени подробно или в по-обобщен вид. Ние се придържаме към формулировка на етапите направени от Косала и Блокийл, които са следните (фиг. 6.2):

1. Откриване на ресурси - търсене и извличане на информация от мрежата. Използват се търсещите машини, които обхождат милиони документи и индексират думите в уеб документите, запазвайки си индексите за последващо използване в други търсения.
2. Извличане на информация - автоматично избиране и предварителна обработка на специфична информация от определени уеб източници. Тази подготовка включва: обработка на лог файлове, преобразуване на текстове, чрез премахване на излишните думи, извличане на фрази и словосъчетания, съставяне на релации. Като резултат от обработката на този етап трябва да се получат данни, които имат подходяща структура за анализ.



Фиг. 6.2. Етапи на процеса на WM

1. Откриване на общите модели на данните. Избира се метод за откриването на знания. За целта се прилагат специални алгоритми за откриване връзки и зависимости в данните. Необходимо е да се отбележи, че на този етап може да се извърши допълнителна обработка на данните, за да се приложат избраните методи.
2. Прилагане на модела и анализиране на получените резултати. На този етап изключително значение има ролята на човека и неговата интерпретация на получените резултати.

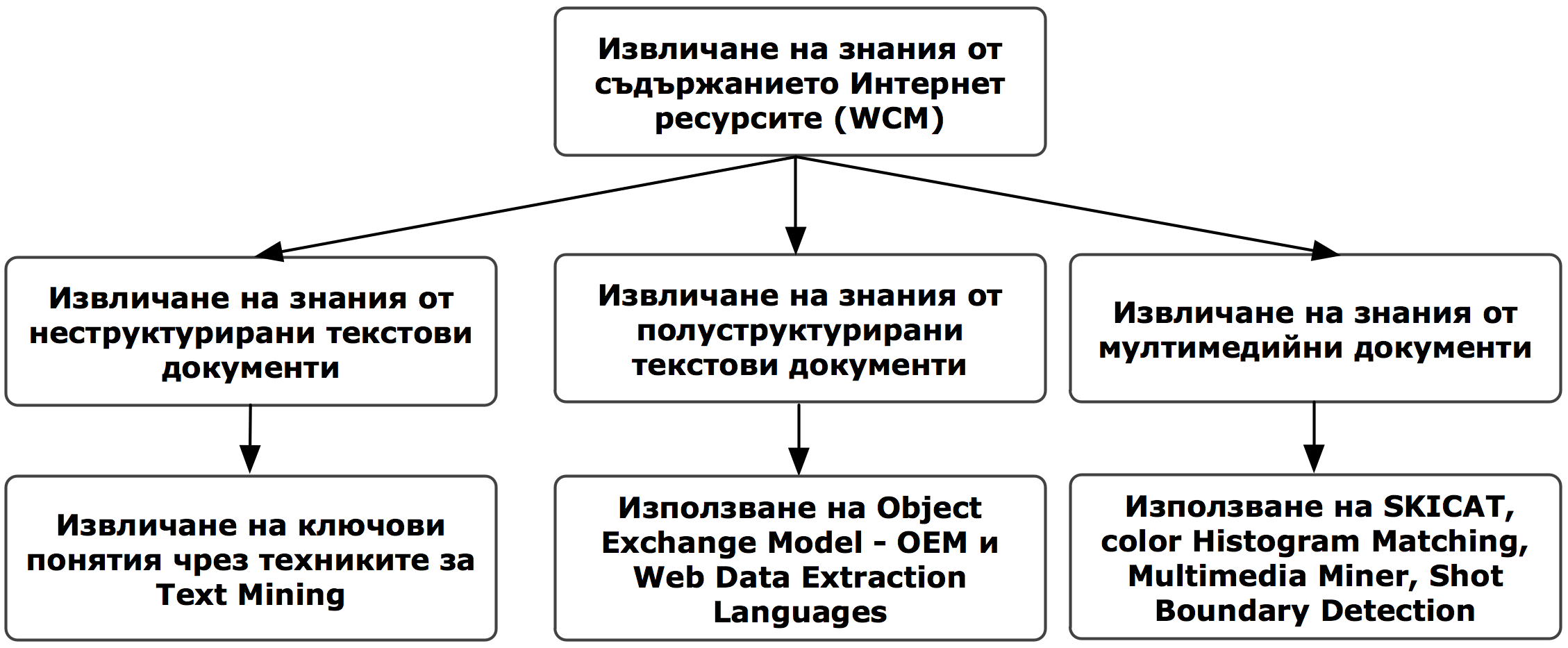
Web Mining включва прилагане на Data Minimg техниките, които са взаимствани от различни дисциплини, като бази от данни, статистика, машинно обучение, разпознаване на шаблони, визуализация, извличане на информация, обработка на образи и сигнали и анализ на пространствени и времеви данни.

**Решаваните задачи** чрез WM са сходни с тези на DM.

**WEB CONTENT MINING - WCM**

# Същност на процеса на извличане на знания от съдържанието на уеб източници

Процесът на извличане на знания от съдържанието на уеб източниците се определя като **Web Content Mining** (WCM). Техниките, които се използват за извличане на информация от съдържанието на документите до голяма степен зависят от типа на ресурсите (фиг. 6.3).



Фиг. 6.3. Техники за извличане на знания от съдържанието на Интернет източници

Най-общо може да се дефинират три групи уеб документи: неструктурирани текстови документи, полуструктурирани текстови документи и мултимедийни документи. Към първата група се отнасят всички текстови документи, които нямат определена структура. Напр. E.mail. Към втората група се причисляват документи, които като цяло имат структура, но съдържат и неструктурирани елементи, текст, такива са HTML, XML документите. Към третата група спадат документите, които съдържат мултимедия, за тях се използват специфични техники, които автоматично се откриват границите между кадрите във видео, връзката между цветови компоненти, класификация на изображения и др.

Търсенето на знания в съдържанието на документите е нетрадиционен процес на откриване на нови, потенциално полезни и разбираеми модели в данните. Процесът на анализ на съдържанието на уеб документите преминава през следните стъпки:

1. Търсене на документи, които да се подложат на анализ. На този етап се използват търсещите машини (търсачки), които позволяват намиране на информация по зададени от потребителя термини (ключови думи). Може да се използват и специално разработени програми за търсене, които работят на базата на алгоритмите на някои от най-известните търсещи машини.
2. Предварителна обработка на документите. В текстовите документи има много думи и не е целесъобразно всички те да се подлагат на анализ, тъй като не всички са носители на полезна информация. Освен това някои от думите са синоними, имат едно и също значение и е добре да се разглеждат като един и същ термин. На този етап се премахват излишните думи, които са спомагателни, носят малко информация за съдържанието на документа, като например: “за”, “по-горе”, “през”, “след”, “така”, “по-късно”, “отново”, “срещу” и др. Всички думи се преобразуват да бъдат изписани с малки букви. Освен това се прави представяне на думите в нормална форма, като се премахва множественото число, наставките и др. Kaто алтернатива преобразуването може да използва представянето на думата като множество от предварително определен брой символи. Например думата “маса” може да се представи като комбинации от 3 символа “\_са”, “мас”, “аса”, “са\_”, като подчертаващото тире е предшествания символ или разделителя след думата. При този способ има по-малко граматически грешки, но се получават и много безсмислени думи. В процеса на предварителната обработка се осъществява и преобразуване на символите, за да бъдат изписани само с малки букви.

3. Извличане на информация. На този етап се използват методите за класификация, асоциация, клъстеризиране и др. Документите се разпределят в категории, извършва се автоматично анотиране, извлича се най-съществената част от текста, като ключови понятия, имена, идентифицират се асоциативните отношения. Откриват се зависимости и модели, които се прилагат за анализ на информацията.

4. Интерпретация на резултатите. На последния етап се визуализират резултатите, обикновено в графичен вид, което позволява да се извърши анализ и тълкуване.

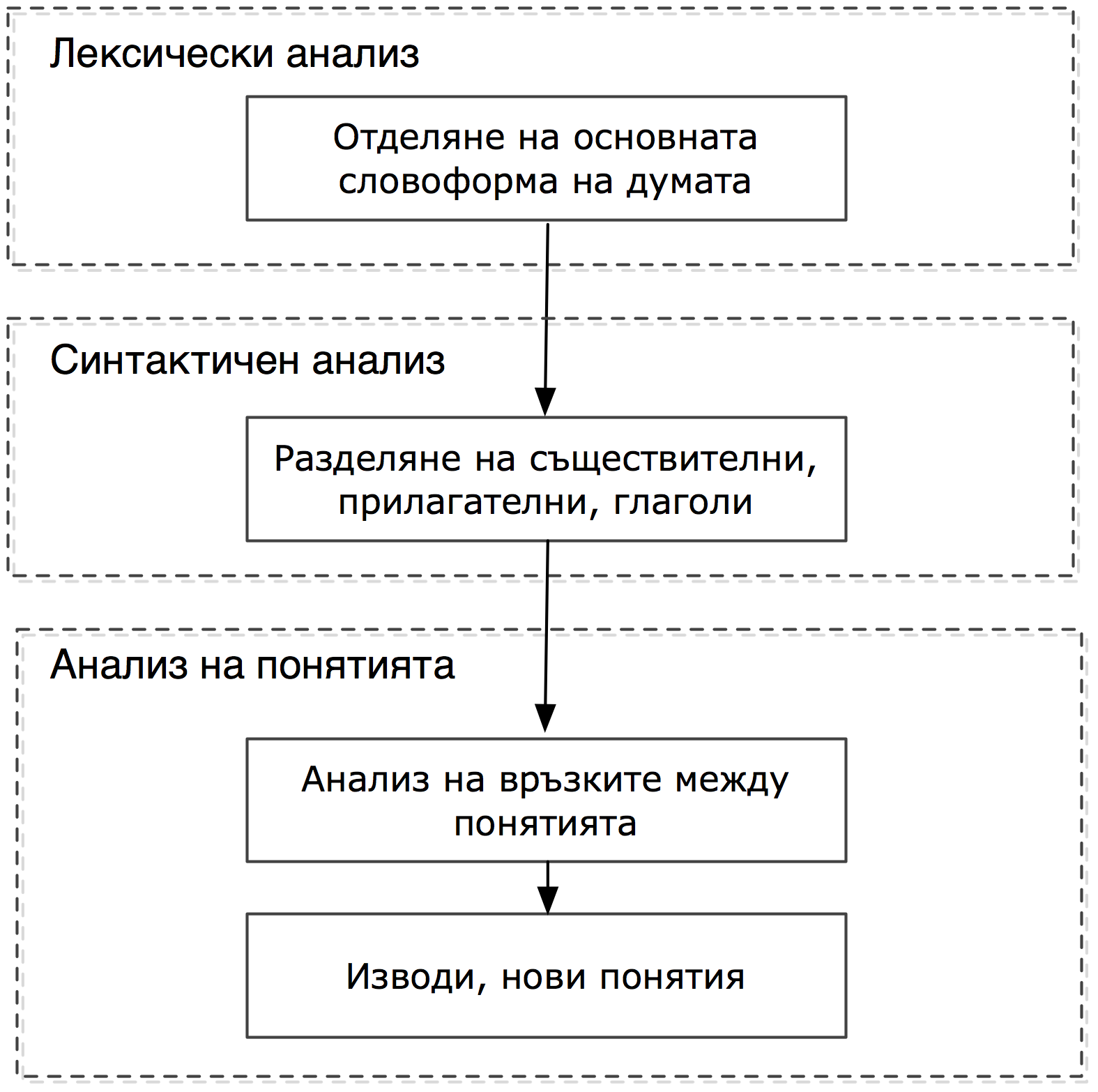
# 6.2. Извличане на ключови понятия от текстови документи (Text Mining)

В основата на Web Content Mining е процесът на извличането на ключови понятия от текстови документи (Text Mining). Този процес може да се разглежда като филтриране на текста и от своя страна се разделя на следните етапи: **лексически анализ, синтактичен анализ и анализ на понятията** (фиг. 6.4).

Лексическият анализ представлява определяне на лексикалните единици (основните форми на думите) в текста. Известно е, че отделните словоформи се образуват чрез граматически правила. В [българския език](http://bg.wikipedia.org/wiki/%D0%91%D1%8A%D0%BB%D0%B3%D0%B0%D1%80%D1%81%D0%BA%D0%B8_%D0%B5%D0%B7%D0%B8%D0%BA) за основна словоформа се избира думата, която е в единствено число, не членувана ако е съществително или прилагателно и в сегашно време, първо лице, единствено число ако словоформата е глагол.

Синтактичният анализ разграничава думите не като лексикални единици, а като части на речта. Определят се главните и второстепенни части в изреченията и думите се разделят на съществителни, прилагателни, глаголи.

Анализът на понятията се осъществява чрез използване на различни методи за клъстеризация, класификация, асоциация, които спомагат за откриване на зависимости между наборите от думи.



Фиг. 6.4. Процес на извличане на ключови понятия

# Алгоритми за извличане на знание от съдържанието на уеб документи

За извличането на знания от съдържанието на уеб документите се използват различни алгоритми и информацията се представя по различен начин в зависимост от това какви точно са текстовите документи. За работа напълно неструктурирани документи се използват алгоритми, чрез които се извличат ключови понятия от текст. Примери за такива алгоритми са дадени в таблица 6.1.

Таблица 6.1.

Алгоритми за извличане на съдържание от неструктурирани документи

|  |  |  |  |
| --- | --- | --- | --- |
| **Автор** | **Алгоритъм и представяне на моделите** | **Начин за представяне на информацията** | **Приложение** |
| Ahonen и др. | Еposode rules -асоциативни правила | колекция от думи | * за намиране на ключови думи и фрази * за откриване на граматически правила и словосъчетания |
| Billsus и Pazzani | * TFIDF - за определяне значе-нието на думите в даден текст   - Naive Bayes - базиран на теоремата на Бейз | колекция от думи | текстова класификация |
| Cohen | * Propositional rule based system - базира се на допускания * Индуктивно логическо   програмиране | релация | текстова класификация |
| Dumais и др. | * TFIDF - за определяне значението на думите в даден текст * Дървета на решения * Naive Bayes - базиран на теоремата на Бейз * Мрежи на Бейс * Support Vector Machines | * колекция от думи * фрази | текстова категоризация |
| Feldman и Dagan | Относителна ентропия | концептуални категории | намиране на модели между текстовите данни |
| Feldman и др. | Асоциативни правила | терми | намиране на модели сред терми в текстовите данни |
| Frank и др. | Naive Bayes - базиран на теоремата на Бейз | фрази | намиране на фрази |
| Freitag и McCallum | Скрити модели на Марков | колекция от думи | обучаващи се модели за извличане |
| Hofmann | Статистически методи | колекция от думи | йерархично групиране |
| Honkela и др. | Self-Organizing Maps - самоорганизиращи се карти | колекция от думи | клъстеризация на текстове и документи |
| Junker и др. | Индуктивно логическо програмиране | релация | * категоризация на текст * обучаващи се модели за извличане |
| Kargupta и др. | * Неконтролируема   йерархична класификация   * Дървета на решения * Статистически методи | колекция от думи | * класификация на текст * йерархично групиране |
| Nahm и Mooney | Дървета на решения | колекция от думи | * предсказване на думи * зависимости |
| Nigam и др. | Максимална ентропия | колекция от думи | текстова класификация |
| Scott и Matwin | Rule based system - базира се на правила | * колекция от думи * фрази * синоними, хиперними9 | текстова класификация |
| Soderland | Обучение по правла | * изречения * клузи | изучаване на правилата на изречението |
| Weiss и др. | Ускорени дървета на решения | колекция от думи | текстова категоризация |
| Wiener и др. | * Невронни мрежи * Логистична регресия | колекция от думи | текстова категоризация |
| Witten и др. | Компресиране на текст | - именувани единици | класификация и наименуване на същности |
| Yang и др. | * алгоритми за клъстеризация * на най-близкия съсед * дървета на решения | * колекция от думи * фрази | откриване на събития и проследяване |

При анализ на полуструктурирани документи се използват по-развити средства за диференциране на текста, тъй като тези документи имат някаква структура. Алгоритмите за анализ използват HTML структурата на уеб страниците и основните методи за извличане на знания като класификация, клъстеризация, асоциация. Основните алгоритми са обобщени и дадени в таблица 6.2.

Таблица 6.2.

Алгоритми за извличане на съдържание от полуструктурирани документи

|  |  |  |  |
| --- | --- | --- | --- |
| **Автор** | **Алгоритъм и представяне на моделите** | **Начин за представяне на информацията** | **Приложение** |
| Craven и др. | * Modified Naive Bayes - модифициран алгоритъм на Бейз * Индуктивно логическо програмиране | релация и онтология | * хипртекстова класификация * извличане на връзки между уеб страници * изучаване на правилата за извличане |
| Crimmins и др. | Unsupervised and supervised classification algorithms - класификационни алгоритми | фрази, URL, метаинформация | * йерархична и графична класификация * клъстеризация |
| Fürnkranz | Обучение по правила | колекция от думи и информационни хипервръзки | - хипертекстова класификация |
| Joachims и др. | * TFIDF - за определяне значението на думите в даден текст * подпомагащо обучение | колекция от думи и информационни хипервръзки | - хипертекстово предсказване |
| Muslea и др. | Обучение по правила | колекция от думи, тагов и позиции на думите | изучаване на правилата за извличане |
| Shavlik и Eliassi-Rad | Невронни мрежи с подсилено обучение | колекция от думи и взаимоотношения | - хипертекстова класификация |
| Singh и др. | * Променени асоциативни правила * Класификационни алгоритми | понятия и наименования на обекти | намиране на шаблони в полуструктуриран текст |
| Soderland | Обучение по правила | изречения, фрази и наименование на обекти | изучаване на правилата за извличане |

Чрез WCM може: автоматично да се класифицират и групират (клъстеризират) уеб страници, според техните теми; да се изследва какво търсят потребителите в Интернет и да се откриват характерни шаблони за търсене на зададената група от хора; да се откриват модели в уеб страниците на базата на описанията на продуктите, съобщения с мнения от форуми и др. Освен това WCM се използва за подобряване намирането и филтрирането на информация, категоризиране на документи, намиране на подобни страници върху различни сървъри, идентифициране на теми, представени в различни уеб документи.

# Извличане на знания от структурата на уеб източници (Web Structure Mining – WSM)

Задачите на извличането на знания чрез изследване на структурата на интернет ресурсите са: да се открие модел на връзките в мрежата и да се изследва структурата на уеб документите (фиг. 6.5). Web Structure Mining се базира на топологията на хипервръзките с или без описание на тези връзки. Откритият модел може да се използва за категоризиране на уеб страници и е полезен при генерирането на информация, например за сходството и отношенията между уеб сайтове.



Фиг.6.5. Извличане на знания от структурата на Интернет ресурси

През последните години структурата на уеб страниците се използва за анализ на важна информация.

Най-общо връзките са 2 вида: in-links - връзки, сочещи към страница, т.н входящи връзки и out-links - връзки, сочещи към други страници, т.н. изходящи връзки. Обикновено колкото е по-голям броят на връзките към дадена страница, за толкова по-полезна се счита тя. Освен това има и алгоритми за анализ, които изследват и текста, който описва връзките (т.н. anchor текст) и може да спомогне за по-качествен анализ.

Най-разпространените WSM алгоритми са: Page Rank и HITS (Hyperlink-Induced Topic Search). Чрез Page Rank се изпълнява класиране на страници с хипертекстови връзки между тях, използван в търсачката Google. Алгоритъмът е написан от Лари Пейдж (откъдето идва и името Пейдж-Ранг) и Сергей Брин.

Основната идея на алгоритъма Page Rank е следната, ако например на една страница а има N връзки към други страници, между които е и страницата b, за потребителя посетил страницата a, вероятността да посети b е 1/N. Определянето на ценността на страниците, на базата на изходящите връзки е количеството престиж, което страницата b получава от страницата а като 1/N oт престижа й. Уеб страницата е с висока оценка, ако е свързана към много други страници и те са с добър рейтинг.

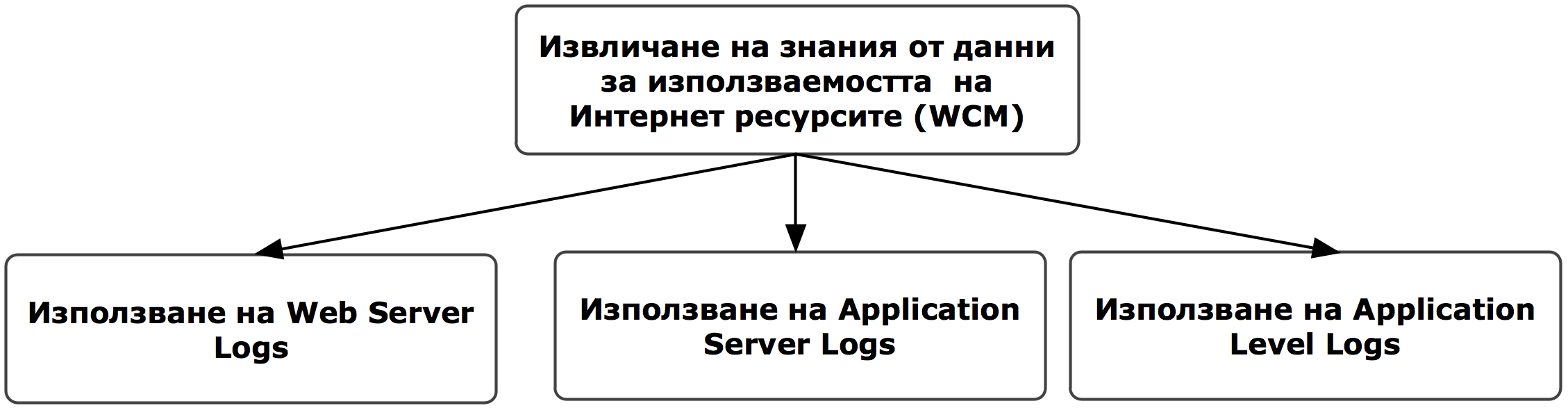
Алгоритъмът HITS, известен още като Hubs and authorities се използва за анализ на връзки. Разработен е от Джон Клайнбърг. Основната идея при него е първо да се фокусира на подходящите страници и след това да се изчисли значимостта им. В сравнение с Page Rank, се работи с много по-малък граф. Започва се с търсене по ключова дума и след това се прави анализ на структурата на връзките. Процесът се нарича още извеждане на “същината на темата”. При този алгоритъм се изчисляват две оценки, за хъб - страница, която сочи към много други и за значимост - страница към която сочат много други страници. В обобщение - страница с висок рейтинг е тази от която има много изходящи връзки, както и страница към която сочат много връзки.

Алгоритмите Page Rank и HITS се основават на свързаните страници и на това, че те принадлежат към една и съща тема. Ако обаче се увеличи броят на връзките, бързо се променя и темата. Увеличаването на страниците води до процес, известен като обобщение на темата. Значителното разширяване на набора от връзки увеличава броя на несвързаните страници и затова трябва да се избягва. Друга нежелана ситуация е, когато един тематичен набор от страници сочи към голям набор страници от друга тема и може да се получи отклонение от темата. Проблем може да възникне и при определянето на ранк и при плътно свързаните страници, които са разположени на един и същи сървър или се получава т.н. семейственост. Един лесен подход за избягване на семействеността е назначаването на тежест от 1/k на вътрешните страници на един сайт с общо k страници.

За преодоляване на проблемите с тези алгоритми за анализ е препоръчително да се използва не една система за определяне на значимостта, а да се приложат множество такива и да се претеглят техните стойности за изчисляване на крайния рейтинг на дадена страница.

# 6.5. Извличане на знания от информацията за ползваемостта на уеб източници (Web Usage Mining – WUM)

Процесът на извличане на знания на базата на използването на Интернет източниците се фокусира върху техники, които помагат да се прогнозира поведението на потребителя, докато той взаимодейства с уеб. Базира се на данните за използването на уеб ресурсите, намиращи се в уеб клиента, прокси сървърите и сървърите (фиг. 6.6).



Фиг. 6.6. Извличане на знания от използваемостта на Интернет ресурсите

WUM има за цел да улови и модели на базата на данните за взаимодействието на потребителите с уеб сайта. Тези модели могат да се използват за сегментиране на потребителите, за подобряване на организацията и структурата на сайта, както и за персонализирано обслужване на потребителите. За разлика от WCM и WSM, основният източник на данни за при WUM са сървърните лог файлове, а не самите уеб страници.

Процесът на анализ, може да се осъществи като първо се извличат данни и се вмъкват в релационни таблици или директно да се използват данните от лог файловете като се прилагат специални техники за извличане и анализ.

В научните изследвания процесът на WUM се разглежда в три етапа: предварителна обработка на данните, откриване на правила и модели и анализ на зависимости[[1]](#footnote-1).

1. Предварителна обработка на данните. Осъществява се като се преобразува първичната информация в абстрактни данни, подходящи за по-нататъшните процеси. Основните информационни източници, достъпни за извличане са: лог файловете, описанията на страниците, топологията на сайта, регистрите на потребителите и анкетите.

Може да се каже, че предварителната обработка е най-трудният етап от WUM. Основните проблеми на този етап произтичат от непълнотата на данните, особено ако се следи само по IP адрес. Някои типични проблеми са чрез един IP адрес могат да се осъществяват много сесии, тъй като някои интернет доставчици предоставят набор от прокси сървъри за потребителите си. Възможно е доставчиците да използват средства за анонимност и да предоставят различни IP адреси за заявки от един потребител. Анализите се затрудняват, когато един и същ потребител осъществява достъпа си от различни компютърни устройства или от различни браузъри. За решаването на проблема с идентифицирането на потребителите се използват и анализират и кукитата. След идентифицирането на потребителите се осъществява разделяне на клик потока на сесии. Този процес също понякога e труден за осъществяване, тъй като не е лесно да се установи в кой момент потребителят е напуснал уеб сайта. За разделянето на сесии обикновено се използва интервал от 30 минути. Освен това анализът на сесиите се затруднява и от някои сървъри, които поддържат променливи за всяка сесия и се налага да се осъществява достъп до сървъра на съдържание, за да се разбере каква заявка е зададена и какво се визуализира при потребителите.

В процеса на WUM е необходимо и осъществяване на предварителната обработка на съдържанието. Това е процес на конвертиране на текст, изображения, скриптове, мултимедия във форма, подходяща за извличане на модели от уеб източници. Приложението на технологиите на Data Mining към съдържанието на уеб сайтовете може да се използва като филтър за входа или изхода на алгоритмите за откриване на модели. Например приложението на алгоритмите за класификация може да ограничи откриването на модели само за разглеждани страници, които са за точно определен продукт. Класификацията или клъстеризирането на съдържанието на страниците помага за определяне на целта на посещението. При предварителната обработка текстовете от страниците се представят като вектори от думи. Изображенията и мултимедийните представяния се заместват от ключови думи. Както беше споменато и при описанието на WCM, съдържанието на статичните уеб страници се обработва сравнително лесно в сравнение с динамичните страници, при които има използване на бази от данни, персонализация и се визуализират различни изгледи. За да се извърши обработка на динамичните страници се използват http заявки, програми тип роботи за търсачки, скриптове за достъп до бази от данни.

Етапът на предварителната обработка за WUM изисква и обработка на структурата - хипервръзките между страниците. Структурата може да се извлече и обработи по същия начин като съдържанието. Тук отново по-трудни за обработка са динамичните страници, които имат различна структура при всяка сървърна сесия.

2. Откриване на правила и модели. Това е следващият етап от WUM. При него се прилагат методи и алгоритми от областта на статистиката, извличането на данни, машинното обучение, разпознаването на образи с цел откриването на модели в процеса на уеб анализиране. Най-често използваните техники за извличане на знания за посетителите са статистическите. Чрез различни видове статистически анализи върху сесийните данни може да се намерят средното време за разглеждане на страницата, най-често посещаваните страници, дължината на навигационния път. На базата на анализ на грешките може да се открият опитите за неауторизиран достъп. Получените знания от статистическите анализи са полезни за подобряване на производителността, сигурността или се използват за маркетингови цели.

3.Анализ на моделите. Това е последният етап от WUM. Целта на този процес е да се филтрират намерените модели и правила и да се запазят само тези, които са значими и интересни за изследването. Използва се обикновено графично представяне на данните или различно оцветяване на стойностите, за да се откроят цялостните модели и тенденции в данните.

1. Srivastava, J. et. al. Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data, // http://portal.acm.org/citation.cfm?id=846188 [↑](#footnote-ref-1)