

Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего профессионального образования

Санкт-Петербургский национальный исследовательский университет

информационных технологий, механики и оптики

Факультет Компьютерных технологий и управления
Кафедра ПБКС

Конспект по дисциплине
Теоретические основы КБ

Угрозы ИБ. Методы оценки уязвимостей.

Закон об информации, защите объектов информатизации - определение информации -> сведения вне зависимости от формы представления.

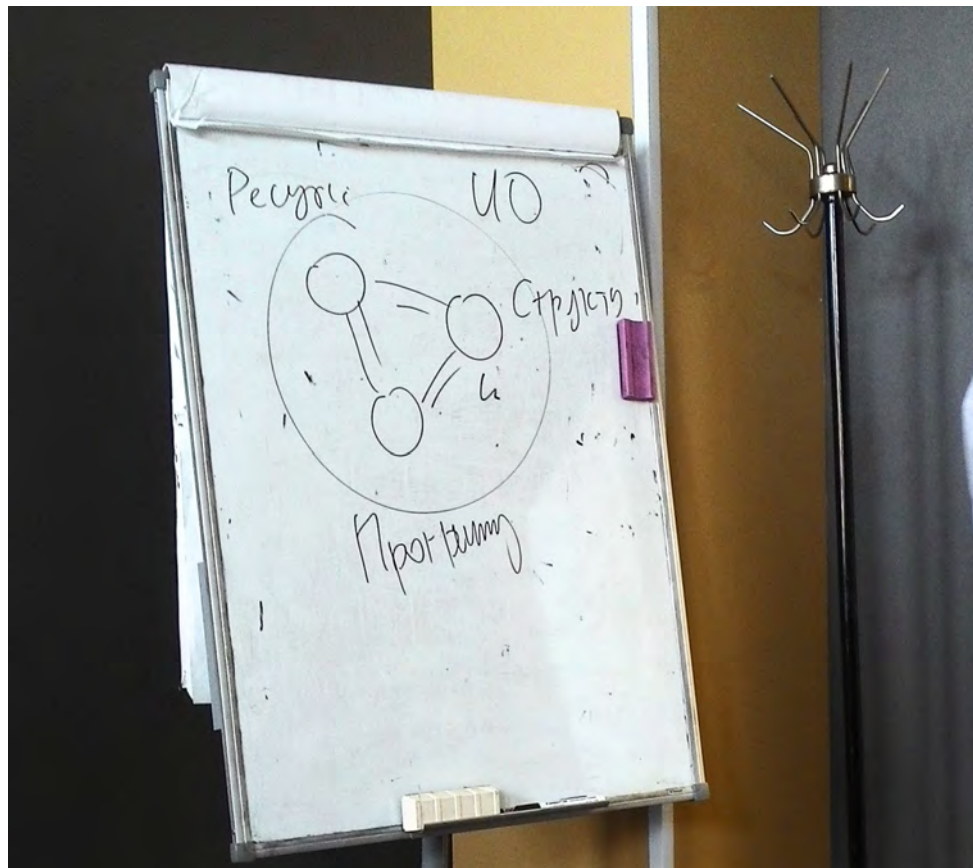
Информация - совокупность данных и соответствующих методов её обработки.

Информация - сведения, необходимые для принятия решения (третий вариант определения).

Сообщение - совокупность зарегистрированных данных.

Информация есть информация, а не материя или энергия.

Информационная сфера - информация, информационная инфраструктура, информационные объекты (носители, потребители), совокупность регламентирующих отношений.



Часто информационные объекты разделяют на критически важные, важные и остальные объекты

Критически важный информационный объект - информационный объект, вывод из строя которого приводит к выходу из строя информационной системы

Важный информационный объект - информационный объект, вывод из строя которого приводит к утрате информационной системой некоторых её функций.

Иной информационный объект - информационный объект, вывод из строя которого не влечёт существенных изменений для информационной системы

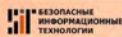
Информационный объект содержит связанные ресурсы, структуры и программы

Вопросы семинара:

1. Общая характеристика ресурса «Банк данных угроз безопасности информации ФСТЭК» <http://bdu.fstec.ru/>
2. Понятие риска ИБ и его способы оценки.
3. Понятие модели угроз. Общая характеристика модели угроз ФСТЭК

Домашние задания

1. В составе учебной группы провести экспертную оценку эффективности программных средств защиты информации на смартфонах (4 произвольных образца СЗИ).
2. Оценить риски:
А) похищения личных данных из домашнего компьютера (смартфона) с использованием злоумышленником удаленного доступа.
Б) похищения личных данных путем кражи злоумышленником персонального компьютера (смартфона).



24

Информационная безопасность - такое состояние рассматриваемой информационной системы, при котором она, с одной стороны, способна противостоять дестабилизирующему воздействию внешних и внутренних информационных угроз, а с другой стороны, её функционирование не создаёт информационных угроз окружающей среде.

Угрозы

Угроза - совокупность факторов и условий, создающих опасность нарушения ИБ организации, вызывающую (или способную вызвать) негативные последствия для организации (ГОСТ 53114-2008)

Базовые угрозы защищаемой информации

Угроза - совокупность факторов и условий, создающих опасность нарушения информационной безопасности организации, вызывающую или способную вызвать негативные последствия (ущерб/вред) для организации. (ГОСТ Р 53114-2008)

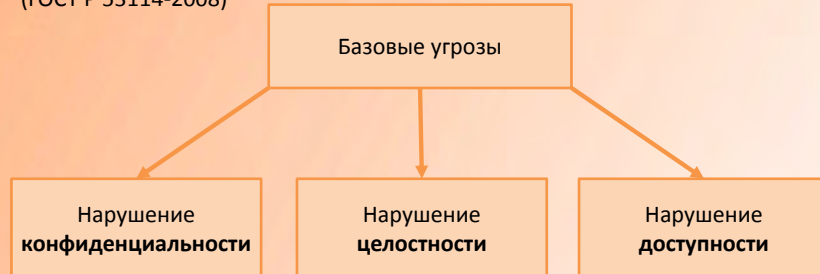
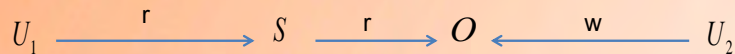


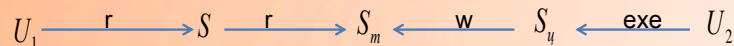
фото - угроза нарушения конфиденциальности

Канал утечки по времени (НСД)

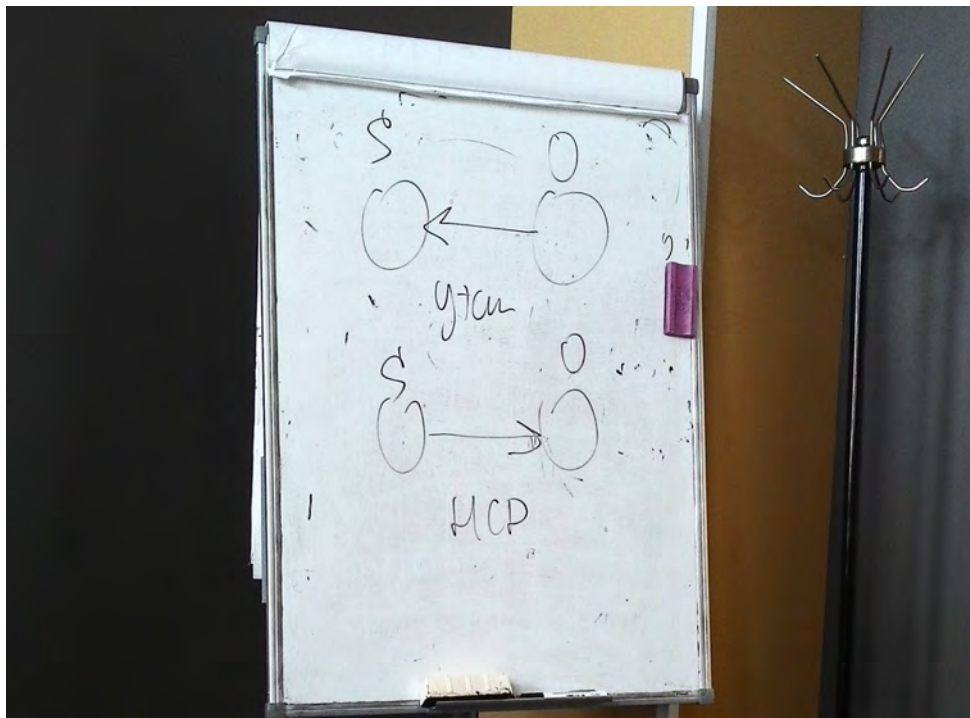
Угроза нарушения конфиденциальности информации. Каналы утечки.



Канал утечки по памяти: пользователь U_1 активизирует процесс, который может получить доступ к общему с пользователем U_2 ресурсу O . При этом U_2 может писать в O , а U_1 читать от S



Канал утечки по времени: U_1 - злоумышленник, U_2 - пользователь, оперирующий ценной информацией, S_y - субъект, информация о котором представляет интерес, S_m - субъект, процесс которого модулируется информацией процесса S_y , S - процесс от имени U_1 , который позволяет наблюдать процесс S_m

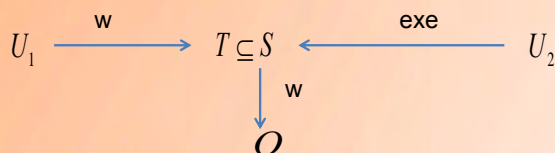


Угроза нарушения целостности информации - незаконные модификация или уничтожение информации.

Угроза нарушения целостности информации

Нарушение целостности информации – это незаконные уничтожение или модификация информации

Канал модификации, использующий «троянского коня»



Злоумышленник U_1 пользуясь правом w модифицировал общий ресурс S , встроив в него скрытую программу T , модифицирующую объект O , при запуске ее пользователем U_2

Классификация угроз по природе происхождения



Каналы НСД



Классификация взято из учебника (забытого =)



Отличие от угрозы в том, что угроза характеризует внешнюю среду, а уязвимость характерна для конкретной системы. **Уязвимость** - свойство, обуславливающее возможность реализации угроз безопасности для обрабатываемой информации.

Оценка угроз и уязвимостей:

прямая экспертная оценка, статистический анализ, факторный анализ

Экспертные оценки

Методы поиска решений не поддающихся формализации задач, основанные на суждениях (оценках) экспертов.

Оценка угроз и уязвимостей



Способы работы с экспертами:

- Интервьюирование
- Анкетирование

Форма выражения оценки может быть явной и неявной

Неявное выражение состоит в том, что эксперт ранжирует оцениваемые элементы (объекты, явления) по степени их важности. Возможен вариант выделения групп с последующим ранжированием объектов внутри группы.

Явная оценка - эксперт даёт элементам лингвистические или количественные оценки (напр, опасно / безопасно / очень опасно). Оценка считается согласованной, если коэффициент конкордации W (коэф. согласованности экспертов) больше 0.75

$$W = \frac{12}{d^2(m^3 - m)} S$$

Экспертные оценки

Экспертными оценками называются такие методы поиска решений сложных, не поддающихся формализации задач, которые основаны на суждениях (оценках, высказываниях) специально выбираемых (назначаемых) экспертов.

Последовательность и содержание решения задач методами экспертных оценок в самом общем виде могут быть представлены следующим образом:

- разработка постановки задачи;
- обоснование перечня и содержания тех параметров задачи, для определения значений которых целесообразно использовать экспертные оценки;
- обоснование форм и способов экспертных оценок;
- разработка реквизитов (бланков, инструкций и т.п.), необходимых для проведения экспертных оценок;
- подбор и подготовка (обучение, инструктаж) экспертов, привлекаемых для решения задачи;
- организация и обеспечение работы экспертов;
- контроль и первичная обработка экспертных оценок;
- базовая обработка экспертных оценок.

В случае, когда в ранжировке есть одинаковые значения - применяют другую формулу (некоторые факторы оценены на одинаковый ранг)

$$W = \frac{12S}{d^2(m^3 - m) - d \sum_{s=1}^d T_s}$$

Оценка согласованности суждений экспертов

d=5 - число экспертов
m=7 - число объектов экспертизы

Номер объекта экспертизы	Оценка эксперта					Сумма рангов	Отклонение от среднего арифметического	Квадрат отклонения от среднего арифметического
	1-го	2-го	3-го	4-го	5-го			
1	4	6	4	4	3	21	1	1
2	3	3	2	3	4	15	-5	25
3	2	2	1	2	2	9	-11	121
4	6	5	6	5	6	28	8	64
5	1	1	3	1	1	7	-13	169
6	5	4	5	6	5	25	5	25
7	7	7	7	7	7	35	15	225
R=						20	S=	630

$$W = \frac{12}{d^2(m^3 - m)} S$$

$$W = \frac{12 \cdot 630}{25 (343 - 7)} = 0,9$$

W= 0,9

Оценка согласованности при наличии связанных рангов

$$W = \frac{12S}{d^2(m^3 - m) - d \sum_{s=1}^d T_s}$$

T_s - показатель связанных рангов в s -й ранжировке

$$T_s = \sum_{k=1}^{H_s} (h_k^3 - h_k)$$

H_s - число групп равных рангов в s -й ранжировке, h_k - число равных рангов в k -й группе связанных рангов при ранжировке s -м экспертом. Если совпадающих рангов нет, то $H_s=0$, $h_k=0$ и, следовательно, $T_s=0$. В этом случае формула (2) совпадает с формулой (1).

Если с согласованностью проблемы - может применяться метод делфи (ака обратная связь) - эксперты знакомятся с анонимными аргументированными обоснованиями суждений коллег.

Оценка привлекательности активов для потенциального злоумышленника

Пример расчета

$d=5$ - число экспертов
 $m=7$ - число объектов экспертизы

Номер объекта экспертизы	Оценка эксперта					Сумма рангов	Отклонение от среднего арифметического	Квадрат отклонения от среднего арифметического
	1-го	2-го	3-го	4-го	5-го			
1	4	4	4	4	3	19	-0,71429	0,510204082
2	3	3	2	3	4	15	-4,71429	22,2244898
3	2	2	1	2	2	9	-10,7143	114,7959184
4	6	5	6	6	6	29	9,285714	86,2244898
5	2	1	3	1	1	8	-11,7143	137,2244898
6	5	4	5	6	6	25	5,285714	27,93877551
7	7	7	6	6	7	33	13,28571	178,5102041
						Rs= 10,71429		Ss= 569,4285714

$h_k=$ 2 2 2 3
 T_s 6 6 6 24
 $\sum T_s$ 42

$$W = \frac{12S}{d^2(m^3 - m) - d \sum_{s=1}^d T_s}$$

W= 0,826467

$$T_s = \sum_{k=1}^{H_s} (h_k^3 - h_k)$$

$$\gamma = \frac{P^U B_0}{C_0}$$

C_0 - стоимость затрат злоумышленника на реализацию угрозы B_0 - профит злоумышленника при реализации угрозы P^U - вероятность реализации угрозы

Оценка привлекательности для злоумышленника информационных активов

Показатель привлекательности угрозы для нарушителя равен:

$$\gamma^U = \frac{P^U B_0}{C_0}$$

Где P^U определяет среднюю меру успеха реализации угрозы.

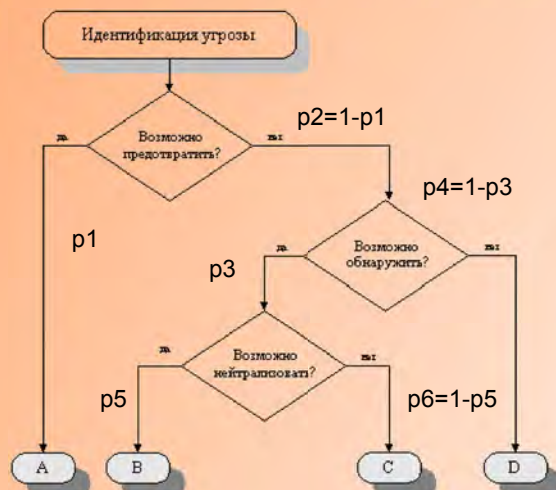
Модель возможных исходов взаимодействия комплекса угроз и СИБ.

Промежуточными в модели могут быть состояния $r_1, r_2, r_3, r_4, r_5, r_6$, характеристики которых отражают следующее:

состояние r_1 - угроза предотвращена, вероятность такого исхода равна p^A
 состояние r_2 - угроза не предотвращена, вероятность такого исхода равна $1 - p^A$
 состояние r_3 - угроза обнаружена, вероятность такого исхода равна p^O
 состояние r_4 - угроза не обнаружена, вероятность такого исхода равна $1 - p^O$
 состояние r_5 - угроза нейтрализована, вероятность такого исхода равна p^H
 состояние r_6 - угроза не нейтрализована, вероятность такого исхода равна $1 - p^H$

Конечная формула

Результирующие события



$$p^A = p^A$$

$$P^B = (1 - p^A) p^O p^H$$

$$P^C = (1 - p^A) p^O (1 - p^H)$$

$$P^D = (1 - p^A) (1 - p^O)$$

Методика оценки уязвимости

Малюк, Герасименко

Привлекательность угрозы

$$P^{AB} = P^H + (1 - P^H)P^O P^H$$

$$P^{CD} = (1 - P^H)((1 - P^H) + P^O(1 - P^H)) = (1 - P^H)(1 - P^O P^H)$$

$$P^U = P^{CD} = (1 - P^H)(1 - P^O P^H)$$

$$\gamma = \frac{(1 - P^H)(1 - P^O P^H)B_0}{C_0}$$

Условия НСД -

- Нарушитель должен получить доступ в контролируемую зону
- Во время нахождения нарушителя в КЗ должен проявиться соответствующий канал НСД
- Нарушитель должен обладать средствами для использования канала НСД
- В канале НСД в момент доступа нарушителя должна быть защищаемая информация.

Эти события случайны.

Условие несанкционированного получения информации

1. Нарушитель должен получить доступ в соответствующую зону.
2. Во время нахождения нарушителя в зоне в ней должен проявиться (иметь место) соответствующий КНСД.
3. Проявившийся КНСД должен быть доступен нарушителю соответствующей категории.
4. В КНСД в момент доступа к нему нарушителя должна находиться защищаемая информация.

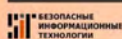
Условные обозначения для формулы оценки уязвимости информации

P_{ikL}^{δ} - вероятность доступа нарушителя k -й категории в L -ю зону i -го компонента системы;

P_{ijL}^k - вероятность наличия (проявления) j -го КНСД в L -й зоне i -го компонента системы;

P_{ijkL}^n - вероятность доступа нарушителя k -й категории к j -му КНСД в L -й зоне i -го компонента при условии доступа нарушителя в зону;

P_{ijL}^u - вероятность наличия защищаемой информации в j -м КНСД в L -й зоне i -го компонента в момент доступа туда нарушителя.



21

L - количество зон; P^{δ} - базовая вероятность реализации угрозы (?)

Пусть K^* есть интересующее нас подмножество из полного множества потенциально возможных нарушителей. Тогда величина $P_{ij\{K^*\}}$ (третья формула) есть ничто иное, как вероятность нарушения защищённости информации указанным подмножеством нарушителей по j -м фактору в i -м компоненте системы

- $P_{ij\{K^*\}}$ Наиболее опасный нарушитель
- $P_{ik\{J^*\}}$ Наиболее небезопасный канал
- $P_{jk\{I^*\}}$ Наиболее уязвимый компонент

REF: DSEC классификация угроз

Оценка уязвимости информации

$$P_{ijkl} = P_{ikL}^{\delta} P_{ijL}^k P_{ijkL}^H P_{ijL}^u$$
$$P_{ijk}^{\delta} = 1 - \prod_{L=1}^5 (1 - P_{ijkl}) = 1 - \prod_{L=1}^5 (1 - P_{ikL}^{\delta} P_{ijL}^k P_{ijkL}^H P_{ijL}^u)$$
$$P_{ij\{K^*\}} = 1 - \prod_{k^*} (1 - P_{ijk}^{\delta})$$
$$P_{ik\{J^*\}} = 1 - \prod_{j^*} (1 - P_{ijk}^{\delta}) \quad P_{jk\{I^*\}} = 1 - \prod_{i^*} (1 - P_{ijk}^{\delta})$$


Общий показатель уязвимости P :

$$P_{\{I^*\}\{J^*\}\{K^*\}} = 1 - \prod_{i^*} (1 - P_{ijk}^{\delta}) \prod_{j^*} (1 - P_{ijk}^{\delta}) \prod_{k^*} (1 - P_{ijk}^{\delta})$$

Риск - вероятность реализации угрозы и нанесение ущерба

Семинар

Оценка риска ИБ

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ		
	НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ	ГОСТ Р ИСО/МЭК 27005- 2010
	Информационная технология МЕТОДЫ И СРЕДСТВА ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ Менеджмент риска информационной безопасности ISO/IEC 27005:2008 Information technology - Security techniques - Information security risk management (IDT)	

Риск информационной безопасности (information security risk):
Возможность того, что данная угроза сможет воспользоваться уязвимостью актива или группы активов и тем самым нанесет ущерб организации.
Примечание - Он измеряется исходя из комбинации вероятности события и его последствия. ($R = P \cdot S$)

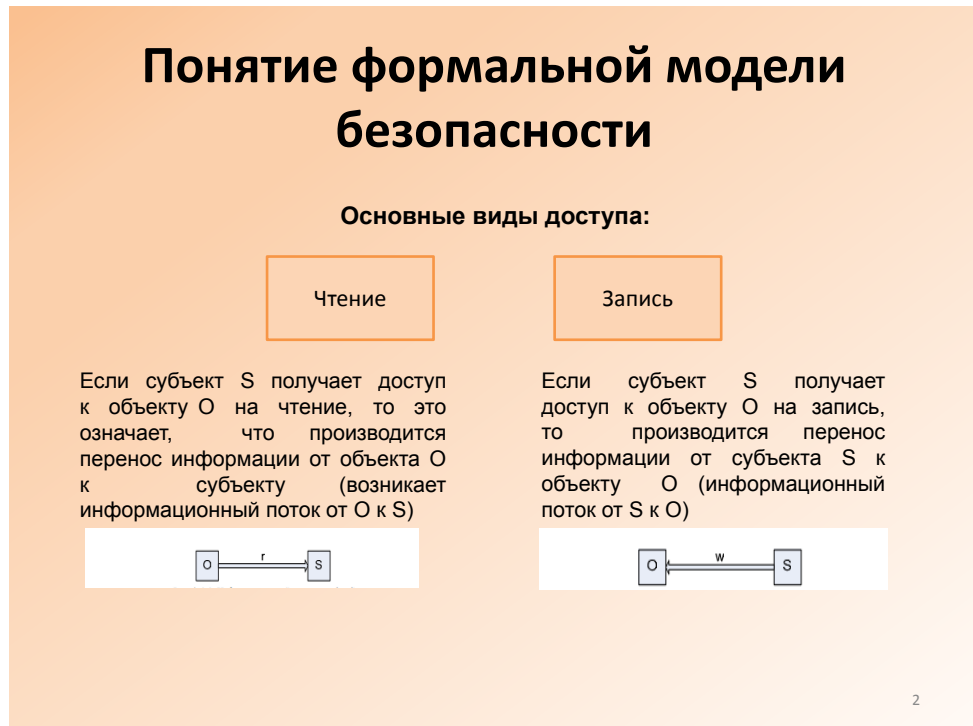
Лекция - через НЕЧ неделю На следующей НЕЧ - вопросы семинара

Пункт 2 - индивидуальный задания на зачёт. 1 половина - ПК 2 половина - Смартфон (по Коновалову) Итог - таким образом видно, что вероятность (угроза является более/менее актуальной)

Время - 15-00

Понятие формальной модели безопасности

Основные виды доступа (на чтение и на запись):



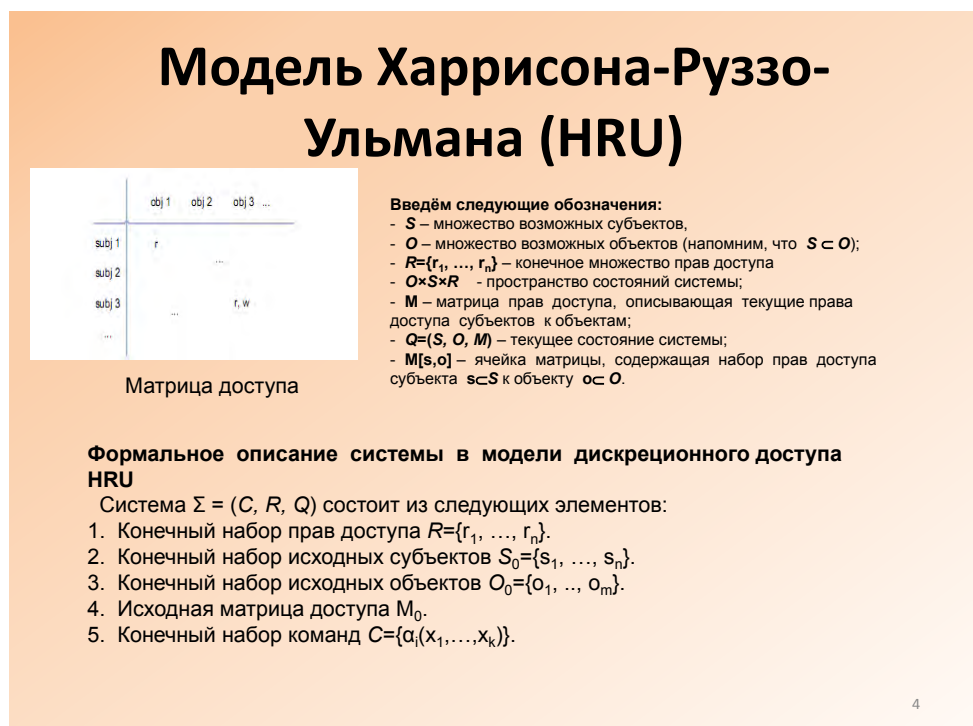
Монитор безопасности обращений



Требования к монитору безопасности обращения

- Ни один запрос на доступ не должен проходить в обход МБО
- Работа монитора безопасности обращений должна быть легко верифицируема

Модель Харрисона-Рузо-Ульмана (HRU)



Множество объектов больше числа субъектов, так как неактивные субъекты попадают в множество объектов.

1-я базовая операция - добавление права

enter r into $M[s,o]$ ($s \in S, \quad o \in O$)

1-я базовая операция – добавление права

1. **enter r into $M[s,o]$** ($s \in S, o \in O$) –
добавление субъекту s права r по отношению к объекту o .
В результате выполнения команды происходят следующие изменения в состоянии системы:
- $S'=S$,
 - $O'=O$,
 - $M'[xs, xo]=M[xs, xo]$, если $(xs, xo) \neq (s,o)$,
 - $M'[s, o]=M[s, o] \cup \{r\}$.

5

Изменяется только выбранная ячейка $[S, O] : M'[s, o] = M[s, o] \cup \{r\}$; остальные ячейки остаются неизменными.

2-я базовая операция - удаление права

delete r from $M[s,o]$ ($s \in S, o \in O$)

2-я базовая операция – удаление права

2. delete r from $M[s,o]$ ($s \in S, o \in O$) – удаление у субъекта s права r по отношению к объекту o .

Изменения в состоянии системы:

- $S' = S$,
- $O' = O$,
- $M'[xs, xo] = M[xs, xo]$, если $(xs, xo) \neq (s, o)$,
- $M'[s, o] = M[s, o] \setminus \{r\}$.

6

Изменяется только выбранная ячейка $[S, O] : M'[s, o] = M[s, o] \setminus \{r\}$; остальные ячейки остаются неизменными.

3-я базовая операция - создание субъекта

Create subject s ($s \notin S$)

3-я базовая операция – создание субъекта

3. create subject s ($s \notin S$) – создание нового субъекта s .

Изменения в состоянии системы:

- $O' = O \cup \{s\}$,
- $S' = S \cup \{s\}$,
- $M'[xs, xo] = M[xs, xo]$ для $\forall (xs, xo) \in S \times O$,
- $M'[s, xo] = \emptyset$ для $\forall ' xo \in O$
- $M'[s, xs] = \emptyset$ для $\forall ' xs \in S$

7

Добавляем по строке/столбцу в множества объектов и субъектов. Все элементы этих строк пусты.

4-я базовая операция - удаление субъекта

`destroy subject s ($s \in S$)`

4-я базовая операция – удаление субъекта

4. **destroy subject s** ($s \in S$) – удаление существующего субъекта s.

Изменения в состоянии системы:

- $S' = S \setminus \{s\}$,
- $O' = O \setminus \{s\}$,
- $M'[xs, xo] = M[xs, xo]$ для $\forall (xs, xo) \in S' \times O'$.

8

Удаляем по строке/столбцу из множества объектов и субъектов.

5-я базовая операция - создание объекта

`Create object o ($o \notin O$)`

5-я базовая операция – создание объекта

5. **create object o** ($o \notin O$) – создание нового объекта o.

Изменения в состоянии системы:

- $O' = O \cup \{o\}$,
- $S' = S$,
- $M'[xs, xo] = M[xs, xo]$ для $\forall (xs, xo) \in S \times O$,
- $M'[xs, o] = \emptyset$ для $\forall xs \in S'$

9

Добавляем по элементу в множества объектов. Все элементы столбца пусты.

6-я базовая операция - удаление объекта

destroy object o ($o \in O$)

6-я базовая операция – удаление объекта

6. **destroy object o** ($o \in O \setminus S$) – удаление существующего объекта o .

Изменения в состоянии системы:

- $O' = O \setminus \{o\}$,
- $S' = S$,
- $M'[x_s, x_o] = M[x_s, x_o]$ для $\forall (x_s, x_o) \in S' \times O'$.

10

Удаляем столбец из множества объектов.

Задачи

Задача 1. Создать субъект, передать ему право владения объектом O .

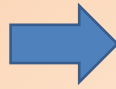
```
1 Program alfn[s,o]
2 create subject s1;
3 enter own into M[s,o];
4
```

Задача 2. Перевести матрицу из состояния 1 в 2;

Задача 1

M

	O1	O2
S1		R,W



M'

	O1	O2	O3
S1		R	W
S2	W	R,W	

11

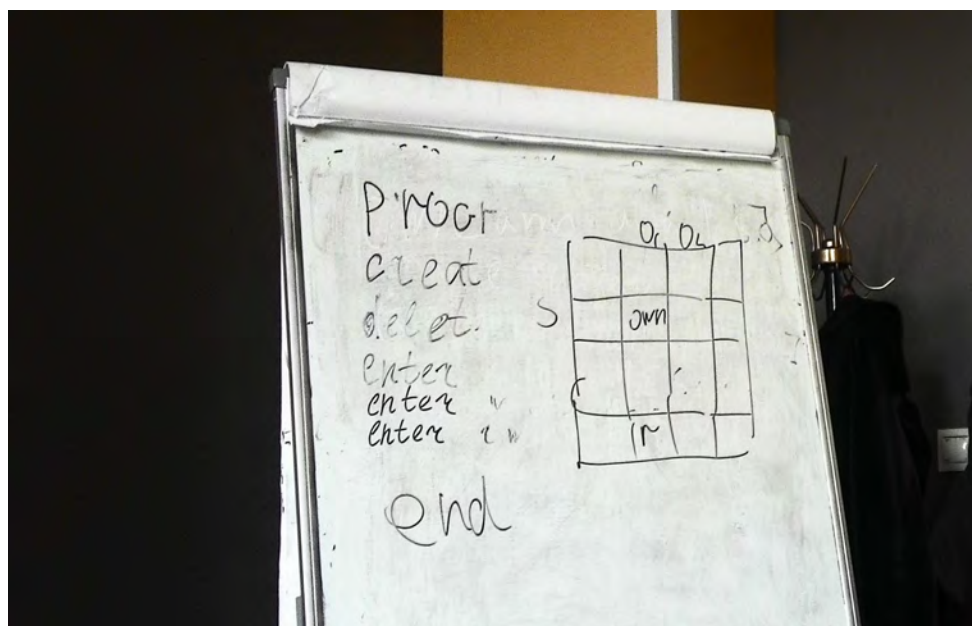
```

1 Program task2[s,o]
2 create subject s2;
3 // – считаем, что создан вместе с субъектом s2 (create object o3);
4 delete W from M[o2,s1];
5 enter W into M[s2,o1];
6 enter R,W into M[s2,o2];
7 enter W into M[s1,o3];
8

```

Задание, которое не будет проверяться: Четырём субъектам присвоено право на владение объектом

Написать программу, которая создаст новый субъект, и добавит право чтения тех объектов, к которым у кого-либо имеется права владения (own)



Модель Take-Grant

Основные положения модели:

1. КС рассматривается как Граф $\Gamma(O, S, E)$, в котором множество вершин представлено:

- множеством объектов O доступа;
- множеством субъектов S доступа ($S \subseteq O$),

Множество ребер представлено множеством E установленных прав доступа (x, y, a) субъекта x к объекту y с правом a из конечного набора прав, в том числе с двумя специфическими правами

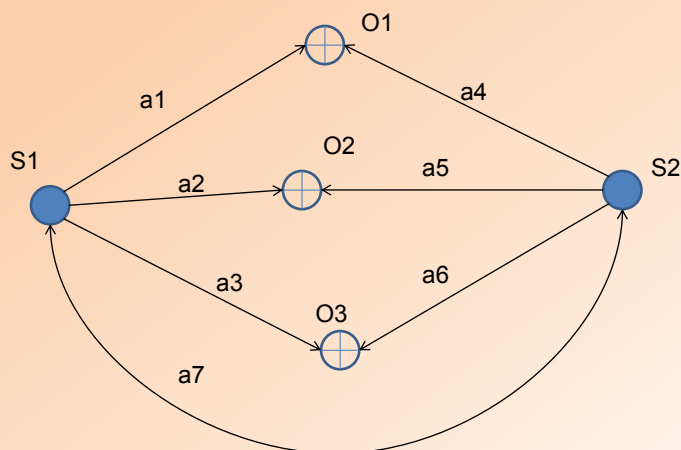
take (t) – право брать права доступа у какого-либо объекта по отношению к другому объекту

grant (g) – право предоставлять права доступа к определенному объекту другому субъекту

12

Модель Take-Grant

Граф доступов



13

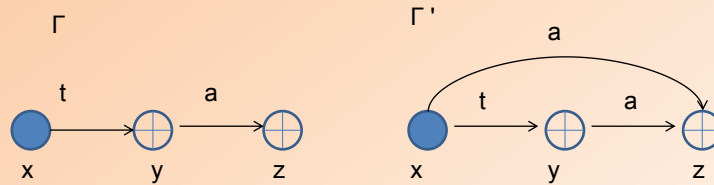
Состояния изменяются под воздействиями 4-х видов

- Команда **take** - «Брать»: $\text{take}(a, x, y, z)$; (право a ; субъектом x ; у субъекта y ; по отношению к объекту z)

Модель Take-Grant

2. Состояния КС изменяются под воздействием **команд** 4-х видов:

2.1. Команда «Брать» - **take(a,x,y,z)**



Изменение состояния фрагмента графа доступов Γ по команде Брать – субъект x берет права доступа a на объект z у объекта y . Переход графа Γ в новое состояние Γ' :

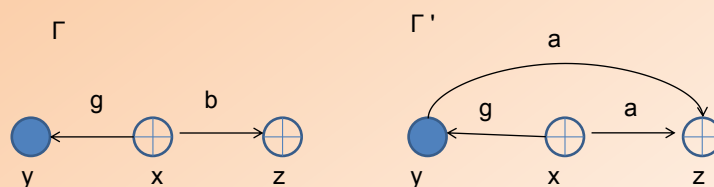
$\vdash_{\text{take}(a,x,y,z)}$

14

- Команда **grant** - «Давать»: $\text{grant}(a,x,y,z)$; (право a ; субъектом x ; у субъекта y ; по отношению к объекту z)

Модель Take-Grant

2.2. Команда «Давать» - **grant(a,x,y,z)**



Субъект x дает объекту y право a на доступ к объекту z

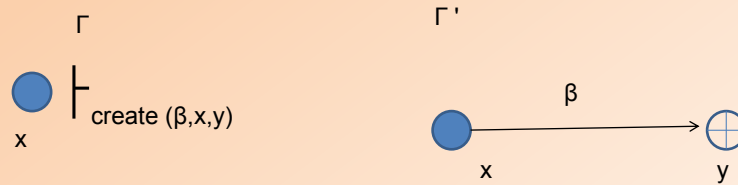
$\vdash_{\text{grant}(a,x,y,z)}$

15

- Команда **create** - «Создать»: $\text{grant}(\beta,x,y)$; (x создаёт объект y с правами доступа на него $\beta \subseteq R$, y - новый объект)

Модель Take-Grant

2.3. Команда «Создать» - $\text{create}(\beta, x, y)$



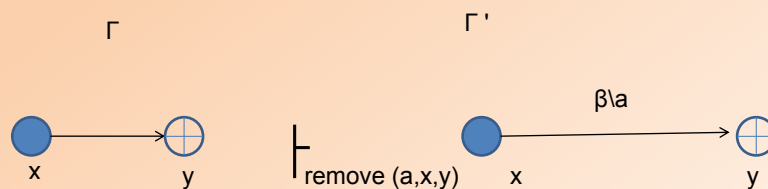
Субъект x создает объект y с правами доступа на него $\beta \subseteq R$, y – новый объект

16

- Команда **remove** - «удалить»: $\text{remove}(a, x, y)$; (x удаляет права доступа на объект y)

Модель Take-Grant

2.4. Команда «Удалить» - $\text{remove}(a, x, y)$



Субъект x удаляет права доступа на объект y

17

Если удалить все права, можем считать субъект удалённым.

Одним из основных вариантов использования модели является анализ на возможность утечки прав доступа.

Модель Take-Grant

3. Безопасность системы рассматривается с точки зрения возможности получения каким-либо субъектом прав доступа к определенному объекту (в начальном состоянии **такие права отсутствуют**) при определенной кооперации субъектов путем последовательного изменения состояния системы на основе выполнения команд.

Предметом анализа при этом являются установленные в начальный момент времени отношения между субъектами по получению и передаче прав доступа на объекты системы, а также возможные ограничения на дальнейшую кооперацию субъектов в процессе функционирования системы.

18

Безопасность системы рассматривается с точки зрения возможности получения субъектом прав доступа к определённым объектам (при этом в начальном состоянии такие права отсутствуют) при определённой кооперации субъектов путём последовательного изменения состояния системы на основе выполнения команд.

Предметом анализа при этом являются установленные в начальный момент времени отношения между субъектами по получению и передаче прав доступа на объекты системы, и возможные ограничения на дальнейшую кооперацию субъектов в процессе функционирования системы.

Не получится ли так, что какой-либо субъект не получит права, которые он получить не должен?

Модель Take-Grant

Команда модели TAKE-GRANT	Условия выполнения	Новое состояние системы
$take(\alpha, x, y, z)$	$x \in S, (x, y, t) \in E, (y, z, \beta)^1 \in E, x \neq z, \alpha \subseteq \beta$	$S'=S, O'=O, E=E' \cup \{(x, z, \alpha)\}$
$grant(\alpha, x, y, z)$	$x \in S, (x, y, g) \in E, (y, z, \beta) \in E, x \neq z, \alpha \subseteq \beta$	$S'=S, O'=O, E=E' \cup \{(y, z, \alpha)\}$
$create(\beta, x, y)$	$x \in S, y \notin O$	$O'=O \cup \{y\}, S'=S \cup \{y\}, \text{ если } y - \text{ субъект } E=E' \cup \{(y, z, \beta)\}$
$remove(\alpha, x, y)$	$x \in S, y \in O, (x, y, \beta) \in E, \alpha \subseteq \beta$	$S'=S, O'=O, E=E' \setminus \{(x, y, \alpha)\} \cup \{(x, y, \beta)\}$

19

Санкционированный доступ

Модель Take-Grant

Санкционированный доступ

Доступ субъекта x к объекту y с правом $\alpha \subseteq R$, отсутствующий в начальном состоянии системы $(x, y, \alpha) \notin E_0$, **возможен тогда и** только тогда, когда существует последовательность перехода системы из состояния в состояние под воздействием команд 2.1, 2.2, 2.3 и 2.4.

Похищение прав

Похищением прав является процесс получения прав доступа на какой-либо объект без предоставления прав третьим субъектам со стороны субъекта, обладающего в начальном состоянии требуемыми правами на объект "интереса".

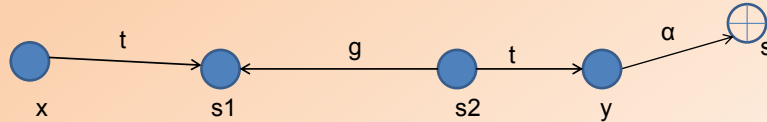
20

Задание на фотке с графом

Модель Take-Grant

Задача

Пусть имеется система субъектов и объектов доступа, представленная Графом доступов $\Gamma_0 (O, S, E)$, в которой сущности x и y связаны tg -путем.



Задание: построить систему команд перехода передачи субъекту x прав доступа α на объект s от субъекта y .

Определение Вершины графа доступов являются tg -связными (со единены tg -путем), если в графе между ними существует такой путь, что каждая дуга этого пути выражает право t или g (без учета направления дуг).

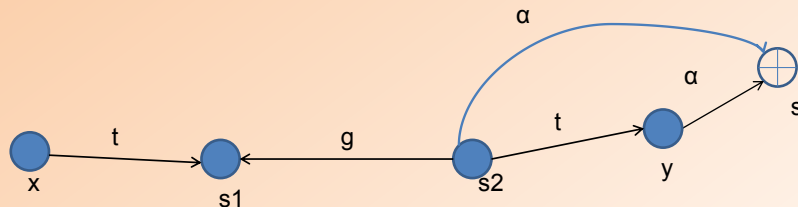
21

Модель Take-Grant

Решение

1-й шаг. Субъект $s2$ на основе своего права t ("брать") на субъект y берет у него право α на объект s – $\vdash takes(\alpha, s2, y, s)$.

Преобразование Графа доступов $\Gamma_0 (O, S, E)$ в новый Граф $\Gamma_1 (O, S, E)$ доступов выглядит следующим образом:

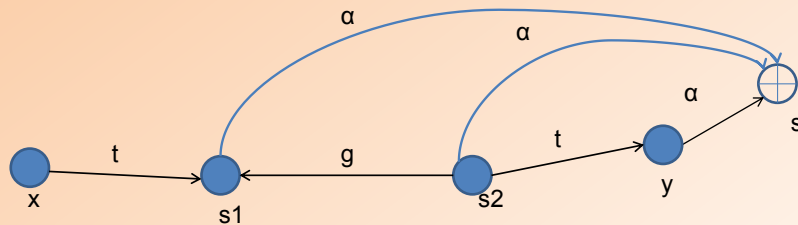


22

Модель Take-Grant

2-й шаг. Субъект $s2$ предоставляет на основе права g ("давать") субъекту $s1$ свое право α на объект s – $\vdash \text{grants}(\alpha, s2, s1, s)$.

Преобразование Графа доступов $\Gamma1 (O, S, E)$ в новый Граф $\Gamma2 (O, S, E)$ доступов выглядит следующим образом

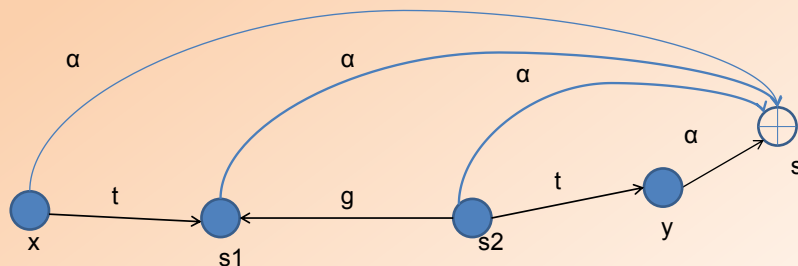


23

Модель Take-Grant

3-й шаг. Субъект x берет на основе своего права t ("брать") на субъект $s1$ имеющееся у него право α на объект s – $\vdash \text{takes}(\alpha, x, s1, s)$.

Преобразование Графа доступов $\Gamma2 (O, S, E)$ в новый Граф $\Gamma3 (O, S, E)$ доступов выглядит следующим образом



24

Управление доступом в распределенных КС

Определение 1. *Распределенной КС называется система, состоящая более чем из одного локального сегмента, представляющего обособленную совокупность субъектов и объектов доступа.*

Основные способы обособления подмножества субъектов и объектов в локальный сегмент:

- **группирование** некоторого подмножества субъектов доступа на основе их порождения и управления одним общим субъектом (системным процессом);
- **локализация** некоторого подмножества субъектов и объектов доступа в рамках некоторой технической компоненты КС;
- **присвоение** всем субъектам и объектам некоторого уникального идентификатора (адреса) в едином информационном (адресном) пространстве и разделение этого пространства на области, обособляющие локальные сегменты.

25

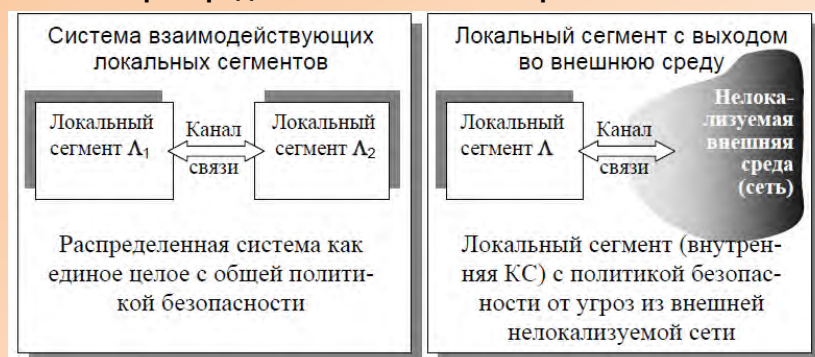
Способы обособления подмножеств субъектов и объектов

- Группирование
- Локализация
- Присвоение идентификатора

Типы распределённых компьютерных систем с точки зрения безопасности

Управление доступом в распределенных КС

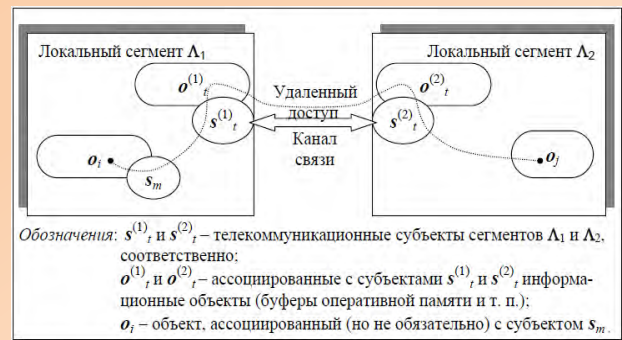
Типы распределенных КС с точки зрения безопасности



Предположение 1. *В локальном сегменте КС функционирует единый для всех субъектов и объектов монитор безопасности, реализующий локальную политику безопасности (политику разграничения доступа).*

26

Структура информационного потока в РКС



Определение 2. Удалённым доступом $p^{out} = \text{Stream}(s_m, o_i) \rightarrow o_j$ субъекта s_m пользователя в локальном сегменте Λ_1 к объекту o_j в локальном сегменте Λ_2 называется порождение субъектом s_m через телекоммуникационные субъекты $s^{(1)}_t$ и $s^{(2)}_t$ локальных сегментов Λ_1 и Λ_2 потока информации между объектом o_j и некоторым(и) объектом o_i сегмента Λ_1 (в т. ч. объект(ы) o_i возможно, но не обязательно ассоциирован(ы) с субъектом s_m).

27

Определение 2 Удалённым доступом $p^{out} = \text{Stream}(S_m, o_i) \rightarrow O_j$ субъекта S_m пользователя в локальном сегменте Λ_1 к объекту o_j в локальном сегменте Λ_2 называется порождение субъектом s_m через телекоммуникационные субъекты $s^{(1)}_t$ и $s^{(2)}_t$ локальных сегментов Λ_1 и Λ_2 потока информации между (далее - см. слайд).

Мы можем говорить о внутризональном разграничении доступа и межзональных правилах.

Управление доступом в распределенных КС

Основные направления обеспечения безопасности в распределенных КС:

- **выделение специального** системного субъекта, обеспечивающего "внешнюю" безопасность;
- **реализация общего**, в том числе возможно с распределенной архитектурой, монитора безопасности, обеспечивающего единую (согласованную) политику безопасности распределенной компьютерной системы

Определение 3. Зоной в распределенной КС называется совокупность подмножества пользователей, подмножества объектов доступа и подмножества физических объектов, обособленных в локальный сегмент с отдельной (внутризональной) политикой безопасности.

28

Определение 3 Зоной в распределённой КС называется совокупность подмножества пользователей, подмножества объектов доступа и подмножества пользователей физи-

ческих объектов, обособленных в локальный сегмент с отдельной (внутризональной) политикой безопасности.

Зональная модель разграничения доступа

Зональная модель разграничения доступа

1. Распределенная КС представляется совокупностью следующих наборов сущностей:

- множество объектов доступа O ;
- множество пользователей U ;
- множество физических объектов системы V (рабочие станции, серверы, коммуникационное оборудование и т.п.);
- множество зон системы Z .

2. Внутризональная политика безопасности в соответствии с предположением 1 реализуется внутризональным монитором безопасности, который обеспечивает весь набор функций безопасности (аутентификация и порождение первичных субъектов доступа пользователей зоны, управление доступом, аудит процессов).

29

Есть внутризональная политика безопасности и межзональная политика безопасности. Внутризональная политика безопасности обеспечивается соответствующим монитором безопасности.

Внутризональный монитор безопасности (определение)

Зональная модель разграничения доступа

Определение 4.

Внутризональным монитором безопасности называется системный субъект (процесс), реализующий в отношении объектов зоны $z \in Z$ **разрешенное множество доступов** которое в общем виде является объединением внутризональных доступов, регламентированных правилами (критериями) внутризональной политики, и удаленных доступов пользователей зоны к объектам других зон, пользователей других зон к объектам данной зоны, разрешенных по правилам (критериям, процедурам) межзональной политики безопасности.

3. На множестве зон системы Z определяется частичный нестрогий порядок, устанавливающий систему межзональных доверительных отношений.

30

Зональная модель разграничения доступа

Элементы системы доверительных отношений:

- **Каждая зона доверяет самой себе** (внутризональные доступы не запрещены).
- **Отношения двухстороннего доверия** (возможны удаленные доступы пользователей первой зоны к объектам второй зоны и наоборот).
- **Отношения одностороннего доверия** (возможны удаленные доступы пользователей одной зоны к объектам другой зоны, но запрещены обратные действия).
- **Отношения доверия между зонами не установлены** (пользователи зон принципиально не доверяют друг другу, доступы не осуществляются).

31

Для того, чтобы получить доступ к внешнему объекту необходимо выполнить вход в зону через межзональную политику, после чего осуществлять запросы через внутризональную политику зоны, в которой запрашиваем объект.

Зональная модель разграничения доступа

4. Процессы доступа пользователей к объектам системы организуются в две фазы:

- "вхождение" пользователя в зону и порождение (первичного) субъекта доступа;
- запрос на доступ субъекта пользователя у внутризонального монитора безопасности к объекту зоны и получение доступа (осуществление потока) в случае удовлетворения запрашиваемого доступа зональной политики безопасности.

Примечание 1. Процедурой вхождения пользователя $u \in U$ в "свою" зону $z \in Z$ называется процесс идентификации/аутентификации и порождения под управлением **внутризонального** монитора безопасности первичного субъекта доступа, осуществляемый пользователем на одной из вычислительных установок зоны.

32

Зональная модель разграничения доступа

*Примечание 2. Процедурой вхождения удаленного пользователя $u \in U$ в "не свою" зону $z' \in Z$ называется процесс идентификации/аутентификации и порождения под управлением **внутризонального монитора** безопасности зоны z' и телекоммуникационных субъектов зон z и z' первичного субъекта удаленного доступа пользователя $u \in U$ в "не своей" зоне z'*

5. Общесистемная политика безопасности складывается из совокупности политики **внутризонального** и **межзонального** разграничения доступа, регламентирующей множество разрешенных (легальных) доступов. При этом процедуры вхождения удаленных пользователей и инициализация в зонах их субъектов производится при выполнении условия:

- ☐ зона вхождения доверяет зоне удаленного пользователя;
- ☐ пользователь уполномочен работать в данной зоне, входящей в подмножество доверенных зон Z .

33

Вопросы к семинару

1. Мандатная модель управления доступом:
 - 1.1. Сущность модели
 - 1.2. Формальное описание
 - 1.3. Примеры использования
2. Ролевая модель управления доступом – сущность и формализация
3. Модель контроля целостности Кларка-Вилсона.

34

10 числа пары нет, след пара - семинар через НЕЧ

След лекция - теория распознавания образов - смотреть теорвер.

Интеллектуальные компьютерные системы

Понятие интеллектуальных компьютерных систем

Важная задача ЗИ: обнаружение угроз без участия человека.

Решение: интеллектуальные (аналитические) системы – концепция, позволяющая компьютерам делать разумные, с точки зрения людей, вещи.

Особенности аналитических систем:

1. Понимание задачи, общего процесса
2. Знание возможностей других систем и людей
3. Связь с пользователем
4. Знания, основанные на здравом смысле
5. Координирование принятия решений и планирование действий
6. Обучение и адаптация

3

Как пример интеллектуальной ИС можно привести эвристический анализатор эвристических систем.

Сигнатурный же анализ может быть представлен такими алгоритмами, как редакционное расстояние (Левенштейна) - применяется в анализаторе DrWeb.

Расстояние Левенштейна

«Дистанция Левенштейна» (LevenShTein diSTance), так же известная как **редакционное расстояние** или **дистанция редактирования**. Это минимальное количество правок одной строки (S) чтобы превратить ее во вторую (T). Под правками подразумеваются три возможные операции:

- стирание символа;
- замена символа;
- вставка символа.

Примеры:

```
levenShTein ('ABC','ABC') = 0
levenShTein('ABC','ABCDEF') = 3
levenShTein('ABC','BCDE') = 3
levenShTein('BCDE','ABCDEF') = 2
```

4

Построение матрицы дистанций Левенштейна

Построение матрицы дистанций начинается из левого верхнего угла матрицы-карты и заканчивается в правом нижнем. Часть матрицы можно заполнить без вычислений: столбец и строка с нулевыми индексами заполняются числами по порядку, начиная с нуля.

Остальные значения дистанций матрицы заполняются по следующим правилам:

1. Если текущие символы строк S и T равны ($S[i-1] = T[j-1]$), то значение ячейки $D[i-1, j-1]$ можно скопировать в ячейку $D[i, j]$. Подобный случай можно сравнивать с движением по карте по диагонали, без штрафа (без затраченных правок). $D[i, j] = D[i-1, j-1]$

5

Построение матрицы дистанций начинается из левого верхнего угла и заканчивается правым нижним.

2. Если же символы $S[i-1] \neq T[j-1]$, то возможны три варианта, из которых выбирается один с минимальной дистанцией:
 - 2.1. Минимальная дистанция слева-вверху - операция замены: символ $S[i-1]$ нужно заменить на $T[j-1]$. $D[i, j] = D[i-1, j-1] + 1$
 - 2.2. Минимальная дистанция слева - операция вставки: символ $T[j-1]$ нужно вставить после $S[i-1]$. $D[i, j] = D[i, j-1] + 1$
 - 2.3. Минимальная дистанция сверху - операция удаления: символ $S[i-1]$ нужно удалить. $D[i, j] = D[i-1, j] + 1$

Пример расчета матрицы

$S='ABC'$ и $T='ABF'$

	A	B	F	
0	1	2	3	
A	1	0	1	2
B	2	1	0	1
C	3	2	1	1

6

Пример расчета матрицы дистанций

Строка 1

$$\begin{bmatrix} & A & B & F \\ 0 & 1 & 2 & 3 \\ A & 1 & ? & ? \\ B & 2 & ? & ? \\ C & 3 & ? & ? \end{bmatrix} \Rightarrow \begin{bmatrix} & A & B & F \\ 0 & 1 & 2 & 3 \\ A & 1 & 0 & 1 \\ B & 2 & ? & ? \\ C & 3 & ? & ? \end{bmatrix}$$

1. A-A: символы совпадают, (п.2) значение берем слева-сверху = 0
2. A-AB: символы различаются, (п.3) слева 0, сверху-слева 1, сверху 2. Берем минимальное значение 0 и прибавляем (0 + 1 = 1). Минимальное было слева, значит операция- вставка. Чтобы A превратить в AB нужно вставить B.
3. A-ABF: A и F различаются, выбираем минимальное из 1, 2 и 3 и прибавляем 1. (1+1= 2) Минимальное значение опять было слева, следовательно операция вставка. Чтобы превратить A в ABF, нужно сначала получить AB (вставка) потом ABF (еще одна вставка)

7

Строка 2

$$\begin{bmatrix} & A & B & F \\ 0 & 1 & 2 & 3 \\ A & 1 & 0 & 1 \\ B & 2 & ? & ? \\ C & 3 & ? & ? \end{bmatrix} \Rightarrow \begin{bmatrix} & A & B & F \\ 0 & 1 & 2 & 3 \\ A & 1 & 0 & 1 \\ B & 2 & 1 & 0 \\ C & 3 & ? & ? \end{bmatrix}$$

1. AB-A: минимальное значение сверху (0), значит операция удаления (0+1=1), итого правок 1. Чтобы из AB получить A нужно удалить B.
2. AB-AB: B и B совпадают, копируем дистанцию слева-сверху (0). Чтобы из AB получить AB никаких правок не нужно
3. AB-ABF: Вставка F.

8

Строка 3

$$\begin{bmatrix} & A & B & F \\ & 0 & 1 & 2 & 3 \\ A & 1 & 0 & 1 & 2 \\ B & 2 & 1 & 0 & 1 \\ C & 3 & ? & ? & ? \end{bmatrix} \Rightarrow \begin{bmatrix} & A & B & F \\ & 0 & 1 & 2 & 3 \\ A & 1 & 0 & 1 & 2 \\ B & 2 & 1 & 0 & 1 \\ C & 3 & 2 & 1 & 1 \end{bmatrix}$$

1. ABC-A: минимальное значение сверху (1), добавляется операция удаления (1+1=2), итого 2 удаления: из ABC нужно удалить BC и получится A.
2. ABC-AB: минимальное значение сверху (0), так как, чтобы ABC превратить в AB, нужно удалить C. Итого 1 правка.
3. ABC-ABF: слева-сверху 0 правок, слева 1, сверху тоже 1 правка. Выбирая наименьшее, мы выполняем замену C на F, что дает результирующее число правок равное 0+1 = 1

Искомая дистанция Левенштейна в этой матрице находится в правом нижнем углу

9

Редакционное предписание - последовательность действий для быстрого преобразования одной строки в другую.

Редакционное предписание

Редакционное предписание — это последовательность действий, необходимых для получения из первой строки второй кратчайшим образом. Обычно действия обозначаются так: D (англ. delete) — удалить, I (англ. insert) — вставить, R (replace) — заменить, M (match) — совпадение.

Для строк ABC и ABF редакционное предписание будет выглядеть так:

M	M	R
A	B	C
A	B	F

10

Интеллектуальные методы можно разделить по методам обучения: «без учителя» и «с учителем»

Общие сведения об алгоритмах кластеризации

Кластеризация – объединение в группы схожих объектов.

Особенности данных:

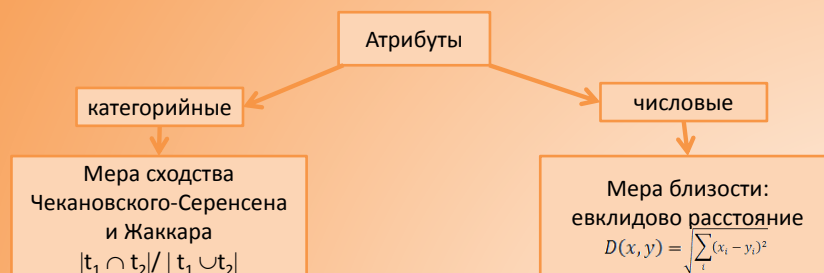
- Высокая размерность
- Большой объем
- Большое количество атрибутов

11

Кластеризация - объединение схожих объектов в группы

Признаки могут быть представлены исчислимыми параметрами - числовыми, остальные - категориальными.

Атрибуты процедуры кластеризации



12

Функция расстояния между кластерами

Функция расстояния одноэлементных кластеров: $R(\{x\}, \{x'\}) = \rho(x, x')$.

Расстояние между новыми кластерами:

$$R(U \cup V, S) = \alpha U R(U, S) + \alpha V R(V, S) + \beta R(U, V) + \gamma |R(U, S) - R(V, S)|.$$

1. Расстояние ближайшего соседа: $R^b(W, S) = \min_{w \in W, s \in S} \rho(w, s)$.

2. Расстояние дальнего соседа: $R^a(W, S) = \max_{w \in W, s \in S} \rho(w, s)$.

13

Иерархическая кластеризация и кластеризация по методу «к-среднего»

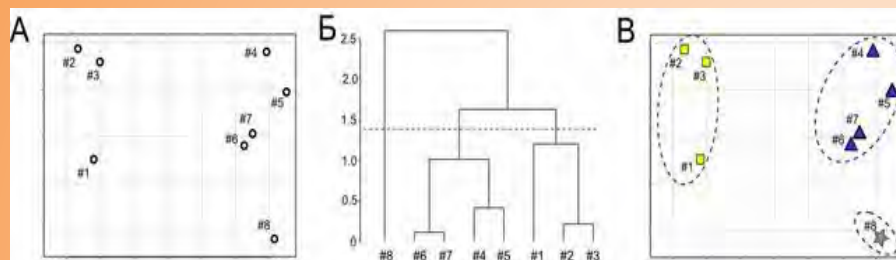
Расстояние между элементами иерархического кластера

3. Среднее расстояние:

$$R^c(W, S) = \frac{1}{|W||S|} \sum_{w \in W} \sum_{s \in S} \rho(w, s)$$

4. Расстояние Уорда:

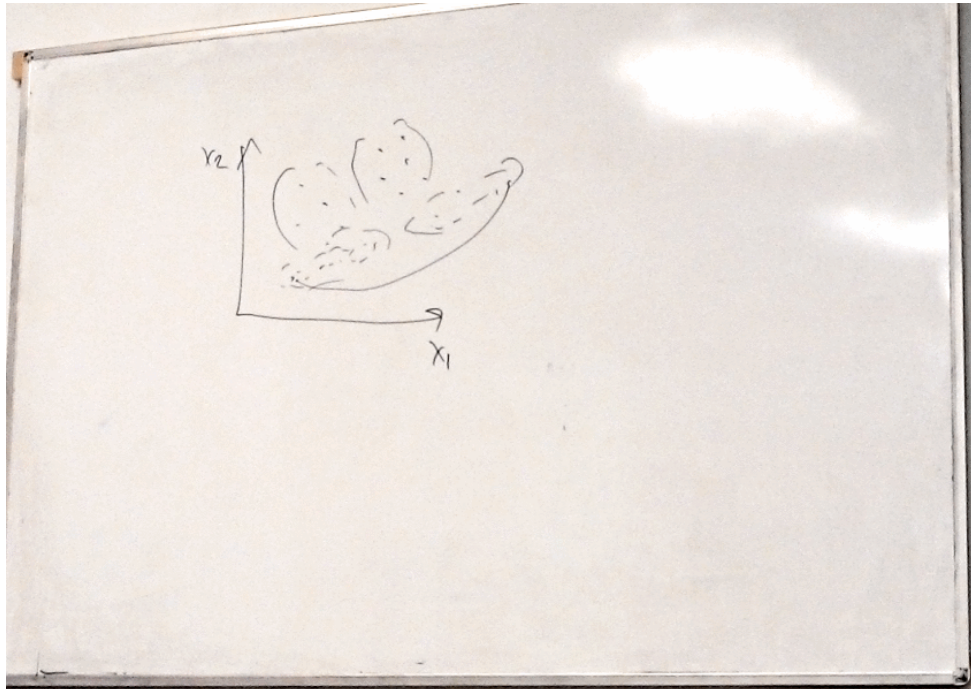
$$R^y(W, S) = \frac{|W||S|}{|W| + |S|} \rho^2 \left(\sum_{w \in W} \frac{w}{|W|}, \sum_{s \in S} \frac{s}{|S|} \right).$$



14

Б - дендрограмма, при построении которых попарно объединяются самые близкие объекты, и объединение продолжается до объединения всех объектов в один кластер. Далее группы делятся в зависимости от желаемого количества кластеров

Если форма кластеров сферическая - используют методы расстояния между ближайшими соседями. Если вытянутая - расстояния между дальними соседями



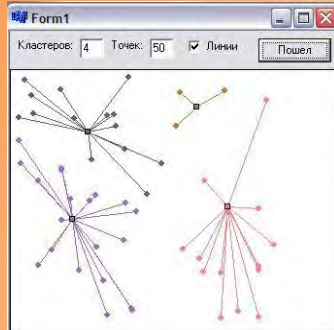
Шаг

- Количество кластеров принимается равным количеству существующих объектов
- два наиболее близких объекта объединяются в кластер
- далее процесс объединения в кластеры повторяется до тех пор, пока все объекты не будут объединены в единственный кластер

Все действия по объединению объектов в кластеры отображаются на дендрограмме, на оси ординат которой отображены расстояния между объединяемыми кластерами объектов, а по оси абсцисс - объединяемые кластеры.

Достоинства иерархического метода заключаются в том, что исследователю нет необходимости заблаговременно определять количество формируемых кластеров.

Алгоритм К-средних (быстрый кластерный анализ)



Функция стоимости критерия суммы квадратов:

$$c(S_i) = \sum_r \sum_s d(x(i,r), x(i,s))$$

где $r = 1..|S_i|$, $s = 1..|S_i|$.

Общий вид алгоритма К-средних:

1. Случайным образом выбираются k центров кластеров - центроидом.
2. Вычисление кластеров в цикле.

2.1. Сначала каждый объект «прикрепляется к тому центроиду, к которому он ближе.

2.2. Когда все точки привязаны, происходит пересчет координат центроидов (см. рис.).

- Случайным образом генерируются 4 центра кластеров.
- Элементы привязываются к центрам на основе евклидова расстояния
- Вычисляются среднее геометрическое полученных кластеров
- На основе позиций новых центров связи пересчитываются, пока ошибка не уменьшится до приемлемого уровня

Если количество кластеров выбрано неверно - надо всё пересчитать заново. Также недостатком является необходимость держать все объекты в памяти.

Алгоритм CLOPE

Назначение: кластеризация очень больших наборов категориальных данных. Достоинства: высокие масштабируемость и скорость работы, а также качество кластеризации, что достигается использованием **глобального критерия оптимизации** на основе **максимизации градиента высоты гистограммы** кластера. Он легко рассчитывается и интерпретируется. Во время работы алгоритм хранит в RAM небольшое количество информации по каждому кластеру и требует минимальное число сканирований набора данных. CLOPE автоматически подбирает количество кластеров, причем это регулируется одним единственным параметром – **коэффициентом отталкивания**.

Основная идея.

Рассмотрим 5 транзакций: $\{(a,b), (a,b,c), (a,c,d), (d,e), (d,e,f)\}$

Сравним качество двух разбиений:

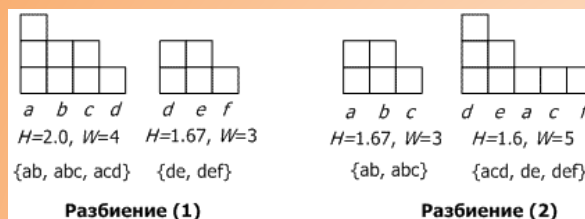
(1) $\{\{ab, abc, acd\}, \{de, def\}\}$

(2) $\{\{ab, abc\}, \{acd, de, def\}\}$

16

Алгоритм CLOPE

Для первого и второго вариантов разбиения в каждом кластере рассчитаем количество вхождений в него каждого элемента транзакции, а затем вычислим площадь (S), ширину (W) и высоту (H) кластера. Например, кластер $\{ab, abc, acd\}$ имеет вхождения $a:3, b:2, c:2$ с $H=2$ и $W=4$.



Разбиение (1) лучше, поскольку обеспечивает большее наложение транзакций друг на друга (соответственно, параметр H там выше).

17

Алгоритм CLOPE

Пусть имеется база транзакций D , состоящая из множества транзакций $\{t_1, t_2, \dots, t_n\}$. Каждая транзакция есть набор объектов $\{i_1, \dots, i_m\}$. Множество кластеров $\{C_1, \dots, C_k\}$ есть разбиение множества $\{t_1, \dots, t_n\}$, такое, что $C_1 \cup \dots \cup C_k = \{t_1, \dots, t_n\}$ и $C_i \cap C_j = \emptyset \forall i \neq j$. Каждый элемент C_i называется кластером, а n , m , k – количество транзакций, количество объектов в базе транзакций и число кластеров соответственно. Каждый кластер C имеет следующие характеристики:

$D(C)$ – множество уникальных объектов;

$Occ(i, C)$ – количество вхождений (частота) объекта i в кластер C ;

$$S(C) = \sum_{i \in D(C)} Occ(i, C) = \sum_{t_i \in C} |t_i|$$

$$W(C) = |D(C)|;$$

$$H(C) = S(C)/W(C).$$

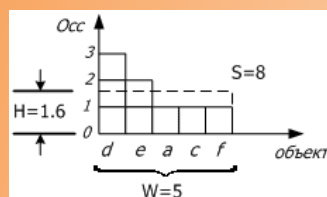
18

На основе базы транзакций (на основе категориальных критериев)

Алгоритм CLOPE

Гистограммой кластера C называется графическое изображение его расчетных характеристик: по оси Ox откладываются объекты кластера в порядке убывания величины $Occ(i, C)$, а сама величина $Occ(i, C)$ – по оси Oy (рис. 2).

Функция стоимости:



$$Profit(C) = \frac{\sum_{i=1}^k G(C_i) * |C_i|}{\sum_{i=1}^k |C_i|} = \frac{\sum_{i=1}^k \frac{S(C_i)}{W(C_i)^r} * |C_i|}{\sum_{i=1}^k |C_i|},$$

где C_i – количество объектов в i -ом кластере, k – количество кластеров, r – коэффициент отталкивания ($0 < r \leq 1$).

С помощью параметра r регулируется уровень сходства транзакций внутри кластера, и, как следствие, финальное количество кластеров. Этот коэффициент подбирается пользователем. Чем больше r , тем ниже уровень сходства и тем больше кластеров будет сгенерировано. Формальная постановка задачи кластеризации алгоритмом CLOPE выглядит следующим образом: для заданных D и r найти разбиение C : $Profit(C, r) \max \rightarrow$.

19

При появлении новой транзакции проверяется лучший вариант изменения функции стоимости, для максимального различия кластеров между собой

Байесовский подход. Основные положения

Проверка гипотез о принадлежности объекта к классу:

$$H_1, H_2, \dots, H_k$$

Априорные распределения вероятности гипотез

$$P(H_i) \quad \text{где} \quad \sum_{i=1}^k P(H_i) = 1$$

Вероятность принадлежности объекта x классу C_i

$$p(H_i/x)$$

20

Пусть имеются K классов объектов

Решение об отнесении объектов к тому или иному классу было принято экспертами. Совокупность объектов, которые система наблюдала ранее и которые были классифицированы экспертами называют обучающей выборкой.

Природа предъявляет наблюдателю (классификатору) новый объект. Необходимо построить решающее правило, согласно которому этот объект будет отнесён к одному из K классов без участия эксперта, при этом решение должно приниматься на основе:

- Характеристик наблюдаемого объекта
- Данных, получаемых в результате анализа обучающей выборки.

Процедура настройки классификатора по данным обучающей выборки называется обучением. Таким образом, необходимо проверить гипотезы $h_1 \dots h_k$ к тому или иному классу.

Обозначим априорные распределения вероятностей гипотез символами $P(H_i)$. Событие принадлежности объекта к тому или иному классу образуют полную группу событий, а значит сумма вероятностей гипотез равна 1. Обозначим вероятность принадлежности объекта x к классу H_i как $P(H_i/x)$.

Основные положения

Если классификатор принимает решение о том, что объект x принадлежит классу C_j , когда на самом деле он принадлежит классу C_i то классификатор несет потери, равные L_{ij} , математическое ожидание которых равно:

$$\pi_j(x) = \sum_{i=1}^k L_{ij} p(H_i | x) \quad \text{Условный средний риск}$$

Апостериорная вероятность принадлежности x классу C_i определяется формулой Байеса

$$p_i = p(H_i | x) = \frac{P(H_i) p(x | H_i)}{p(x)} \quad p(x) = \sum_{i=1}^k p(H_i) p(x | H_i)$$

21

Если классификатор принимает решение о том, что объект $x \in C_j$, а на самом деле $x \in C_i$ - классификатор несет потери L_{ij} , математическое ожидание которого приведено формуле слайда - условный средний риск.

Основные положения

$p(x|H_i)$ есть плотность распределения элементов вектора x при условии, что он принадлежит классу C_i . Величины $p(x|H_i)$ называют функциями правдоподобия x по отношению к H_i .

Тогда условный средний риск равен

$$\pi_j(x) = \frac{1}{p(x)} \sum_{i=1}^k L_{ij} P(H_i) p(x | H_i), \quad j = 1, 2, \dots, k$$

$$\pi_j(x) = \sum_{i=1}^k L_{ij} P(H_i) p(x | H_i), \quad j = 1, 2, \dots, k$$

22

$p(x|H_i)$ - плотность распределения x при условии что он принадлежит классу C_i - функции правдоподобия x по отношению к H_i . Тогда можем посчитать значение условного среднего риска, на основе которого можем вынести решение о принятии гипотезы. Тогда для принятия решения об отнесении предъявленного объекта к одному из классов мы должны вычислить условные средние риски ожидаемых потерь и выбрать среди них ту гипотезу, которая характеризуется наименьшей величиной условного среднего риска.

Двухклассовая классификация

Средний риск при выборе класса C1

$$\pi_1(x) = L_{11}P(H_1)p(x|H_1) + L_{21}P(H_2)p(x|H_2)$$

При выборе класса C2

$$\pi_2(x) = L_{12}P(H_1)p(x|H_1) + L_{22}P(H_2)p(x|H_2)$$

Байесовский классификатор обеспечивает отнесение объекта x к классу с наименьшим значением средних потерь. Поэтому объект x причисляется к классу C1, если выполняется условие

$$\pi_1(x) < \pi_2(x)$$

23

В таком случае можно упростить формулы высчисления среднего риска.

Двухклассовая классификация

Тогда

$$L_{11}P(H_1)p(x|H_1) + L_{21}P(H_2)p(x|H_2) < L_{12}P(H_1)p(x|H_1) + L_{22}P(H_2)p(x|H_2)$$

$$(L_{21} - L_{22})p(x|H_2)P(H_2) < (L_{12} - L_{11})p(x|H_1)P(H_1)$$

Потери от ошибочно принятого решения

выше «потерь» при правильном выборе:

$$L_{ij} > L_{ii}$$

$$\frac{P(H_1)p(x|H_1)}{P(H_2)p(x|H_2)} > \frac{L_{21} - L_{22}}{L_{12} - L_{11}}$$

$$\frac{p(x|H_1)}{p(x|H_2)} > \frac{P(H_2)(L_{21} - L_{22})}{P(H_1)(L_{12} - L_{11})}$$

24

Двухклассовая классификация

$$\Lambda(x) = \frac{p(x/H_1)}{p(x/H_2)} \quad \text{отношение правдоподобия}$$

$$\eta = \frac{P(H_2)(L_{21} - L_{22})}{P(H_1)(L_{12} - L_{11})} \quad \text{пороговое значение}$$

$$x \in C_1, \text{ если } \Lambda(x) > \eta$$

		Верная гипотеза	
		H_0	H_1
Результат применения критерия	H_0	H_0 верно принята	H_0 неверно принята (Ошибка второго рода)
	H_1	H_0 неверно отвергнута (Ошибка первого рода)	H_0 верно отвергнута

25

Ошибки первого и второго рода.

Простейший пример двухклассовой классификации

Наблюдаемая величина – смесь сигнала и помехи: $y = x + n$

Либо только помеха: $y = n$

В общем случае: $y = Ax + n$

Решающая функция, которая в зависимости от реализации y принимает значение 0 или 1.

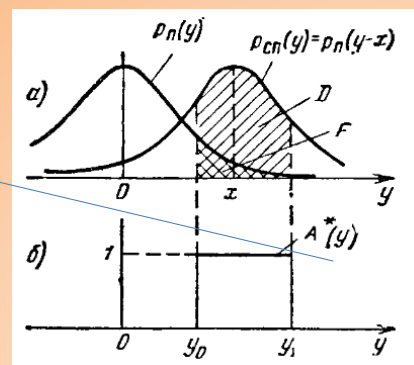
Плотности вероятности случайной величины y при условиях отсутствия сигнала

$A = A_0 = 0$

и при наличии

$A = A_1 = 1$:

$$p(y | A_0) = p_n(y), \quad p(y | A_1) = p_{cn}(y)$$



26

Простейший пример двухклассовой классификации

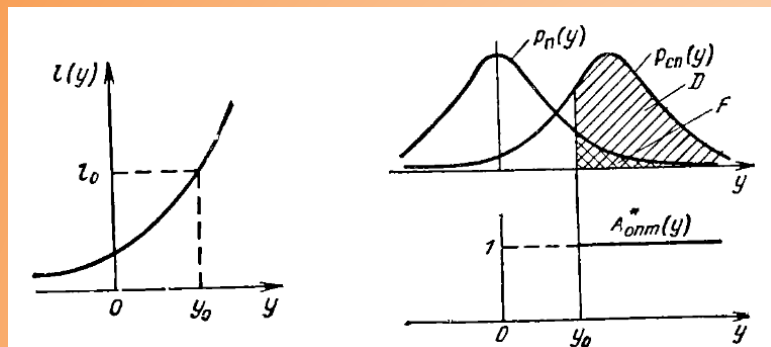
Для случая гауссового распределения сигнала помехи:

$$p_n(y) = \frac{1}{\sqrt{2\pi}n_0} e^{-\frac{y^2}{2n_0^2}} \quad p_{cn}(y) = \frac{1}{\sqrt{2\pi}n_0} e^{-\frac{(y-x)^2}{2n_0^2}}$$

$$l(y) = \frac{e^{-\frac{(y-x)^2}{2n_0^2}}}{e^{-\frac{y^2}{2n_0^2}}} = e^{-\frac{x^2}{2n_0^2}} e^{\frac{xy}{n_0^2}}$$

27

Результат работы классификатора



D – правильное решение (правильное обнаружение)
F – ошибка второго рода (ложная тревога)

28

Фильтрация спама

	Spam						Ham				
1	d	s	g	e		1	u	r	t	o	
2	d	m	f	e		2	w	t	u	j	
3	w	e	g	f		3	k	l	m	b	
4	d	g	r	k		4	r	g	h	o	
5	l	t	o	e		5	r	t	y	u	
6	h	w	e	m		6	h	k	l	e	
7	d	r	u	e		7	l	d	e	e	
						8	k	l	p	m	
						9	s	u	i	o	
						10	p	y	k	l	

Сообщение : o l g m

29

Пример со спамом

Фильтрация спама

$$\Lambda(x) = \frac{p(x/H_1)}{p(x/H_2)} \quad \eta = \frac{P(H_2)(L_{21} - L_{22})}{P(H_1)(L_{12} - L_{11})} \quad x \in C_1, \text{ если } \Lambda(x) > \eta$$

Гипотеза H_1 – сообщение «письмо», гипотеза H_2 – сообщение спам

Пусть $L_{11} = L_{22} = 0$, $L_{12} = 100$, $L_{21} = 10$

Тогда пороговое значение равно:

$$\eta = \frac{7/17 * (10 - 0)}{10/17 * (100 - 0)} = \frac{4,12}{58,8} = 0,07$$

Отношение правдоподобия

(при условии независимости СВ):

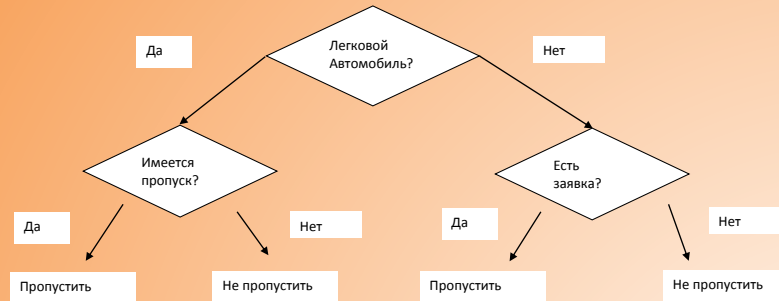
	"o"	"l"	"q"	"m"
$p(x_i/H_1)$	3/40	5/40	1/40	2/40
$p(x_i/H_2)$	1/24	1/24	3/24	2/24

$$\Lambda(x) = \frac{p(x/H_1)}{p(x/H_2)} = \frac{p(x_1/H_1)p(x_2/H_1)p(x_3/H_1)p(x_4/H_1)}{p(x_1/H_2)p(x_2/H_2)p(x_3/H_2)p(x_4/H_2)} = 0,64$$

30

В данном случае считаем вероятности появления каждого из слов независимыми.

Метод деревьев решений



Распространенные алгоритмы: CARP, C4.5 и масштабируемый алгоритм Sprint.

31

Процесс конструирования дерева решений

Пример.

IP	Cookies	Версия ОС	Версия браузер	Результат
Известен	Есть	Без изменений	Без изменений	Класс 1
Известен	Есть	Без изменений	Изменения	Класс 2
Известен	Есть	Изменения	Изменения	Класс 2
Не известен	Есть	Изменения	Изменения	Класс 2
Не известен	Нет	Изменения	Изменения	Класс 1
Не известен	Есть	Изменения	Без изменений	Класс 2
Известен	Нет	Без изменений	Без изменений	Класс 1
Не известен	Нет	Без изменений	Изменения	?

32

Определение 2. Предположим, что множество A элементов, некоторые из которых обладают свойством S , классифицировано посредством атрибута Q , имеющего q возможных значений. Тогда прирост информации (information gain) определяется как

$$Gain(A, Q) = H(A, S) - \sum_{i=1}^q \frac{|A_i|}{|A|} H(A_i, S)$$

где A_i – множество элементов A , на которых атрибут Q имеет значение i .

Решение задачи. Приросты информации для различных атрибутов

$$H(A, \text{Результат}) = -\frac{4}{7} \log_2 \frac{4}{7} - \frac{3}{7} \log_2 \frac{3}{7} \approx 0,985$$

35

Оцениваем прирост информации.

$$\begin{aligned} Gain(A, IP) &= H(A, \text{Результат}) - \frac{4}{7} H(A_{\text{известен}}, \text{Результат}) - \frac{3}{7} H(A_{\text{неизвестен}}, \text{Результат}) \approx \\ &\approx 0.985 - \frac{4}{7} \left(-\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} \right) - \frac{3}{7} \left(-\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3} \right) \approx 0.0202 \\ Gain(A, Cookies) &= H(A, \text{Результат}) - \frac{5}{7} H(A_{\text{есть}}, \text{Результат}) - \frac{2}{7} H(A_{\text{нет}}, \text{Результат}) \approx \\ &\approx 0.985 - \frac{5}{7} \left(-\frac{1}{5} \log_2 \frac{1}{5} - \frac{4}{5} \log_2 \frac{4}{5} \right) - \frac{2}{7} (0) \approx 0.4696 \\ Gain(A, OC) &= H(A, \text{Результат}) - \frac{3}{7} H(A_{\text{безизмен}}, \text{Результат}) - \frac{4}{7} H(A_{\text{измен}}, \text{Результат}) \approx \\ &\approx 0.985 - \frac{3}{7} \left(-\frac{1}{3} \log_2 \frac{1}{3} - \frac{2}{3} \log_2 \frac{2}{3} \right) - \frac{4}{7} \left(-\frac{1}{4} \log_2 \frac{1}{4} - \frac{3}{4} \log_2 \frac{3}{4} \right) \approx 0.1281 \end{aligned}$$

36

дз - построить дерево решений для варианта, когда корнем является + фото

+ дз (индивидуальное) Обучающая выборка состоит из объектов двух классов. Каждый объект характеризуется двумя параметрами : x и y . ($x_1; y_1$ - объект 1 класса, $x_2; y_2$ - второго). Известно, что случайные величины x и y распределены по нормальному закону. Задача - построить байесовский классификатор и отнести распознаваемый объект (последний) к классу 1 или 2.

Иначе задача идеентификации - Дать ответ на вопрос, является

Базовая схема идентификации и аутентификации

1. Ввод идентификатора
2. Проверка актуальности идентификатора (наличие в базе) (нет - следующий список)
3. (да) Аутентификация - ввод секретных для пользователя данных
4. Проверка совпадения аутентификатора (нет - следующий список)
5. (да) Авторизация - получение доступа



Ветвь "Нет"

1. Проверка допустимости повторной попытки (да) - уведомление пользователя, переход в начало исполнения алгоритма
2. (нет) Блокировка пользователя и вывод предупреждения

Всё множество используемых в настоящий момент способов аутентификации можно разделить на 4 группы"

- Методы, основанные на знании некоторой секретной информации (пароли)
- Методы, основанные на использовании уникального предмета (rfid, смарт-карты)
- Методы, основанные на использовании биометрических характеристик человека (отпечаток пальца, рисунок сетчатки, походка, голос, фотография лица, почерк, поведенческие модели)
- Методы, основанные на информации, ассоциированной с пользователем (координаты геолокации)

Особенности парольных систем аутентификации

Является наиболее распространённой. (простота реализации, традиционность)

Методы повышения стойкости

- Увеличение мощности алфавита
- Увеличение длины пароля
- Ограничение скорости перебора
- Отбраковка новых словарных паролей

Основные способы компрометации паролей

1. Использование слабости человеческого фактора (подглядывание, подслушивание, шантаж, угрозы, использование чужих учётных записей с разрешения их законных владельцев)
2. Подбор (полный перебор, использование словарей, использование сведений о пользователе)
3. Использование недостатков в реализации парольной системы (уязвимости приложения и ОС)

Средства контроля периметра

Требования к МЭ из руководящих документов ФСТЭК



Принципы обеспечения целостности информации в компьютерных системах

1. Корректность транзакций - принцип требует обеспечение невозможности произвольной модификации данных пользователя. (Кларк - Вилсон, модель обеспечения целостности)
2. Обеспечение аутентификации пользователей - изменение данных может быть проведено только аутентифицированными пользователями.
3. Минимизация привилегий - пользователи и процессы должны быть наделены теми и только теми привилегиями, которые минимально достаточны для выполнения своих функций.
4. Разделение обязанностей - для выполнения критических или необратимых операций требуется участие нескольких независимых пользователей.

5. Аудит произошедших событий - принцип требует создание механизма подотчётности пользователей, позволяющего отследить моменты нарушения целостности информации.
6. Объективный контроль.
7. Управление передачей привилегий - порядок передачи привилегий должен полностью соответствовать организационной структуре предприятия.
8. Криптографические методы - ЭЦП, хэширование
- 9.
- 10.

Ассиметричная криптосистема



Требования к ЭЦП

Фильтры пакетов

- . Фильтрация пакетов обычно осуществляется по следующим критериям:
 - IP-адрес источника;
 - IP-адрес получателя;
 - порт источника;
 - порт получателя;
 - специфические параметры заголовков сетевых пакетов.
- Фильтрация реализуется путём сравнения перечисленных параметров заголовков сетевых пакетов с базой правил фильтрации.

1. ЭЦП должна доказывать, что подписал документ именно законный владелец и никто другой подписал документ.
2. ЭЦП должна быть неотъемлемой частью документа
3. Невозможность изменения подписанного документа
4. Юридическая доказуемость
5. Невозможность отказа от подписи

Принципы построения систем защиты от угроз нарушения доступности информации

В общем случае обеспечение защиты от угроз нарушения доступности информации реализуется путём создания той или иной избыточности.



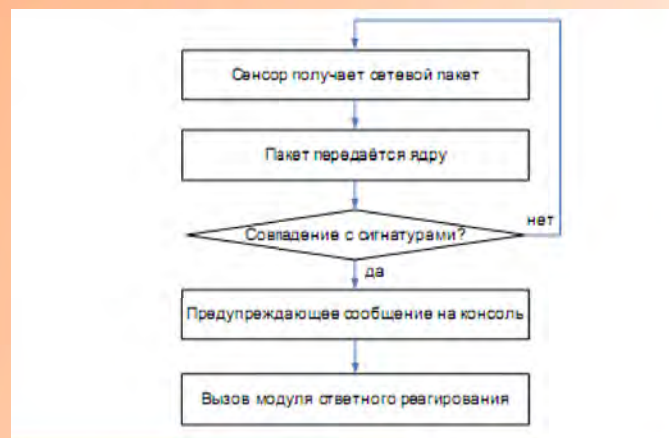
- Дублирование каналов и оборудования
- Запас пропускной способности оборудования
- Кластеризация
- Резервное копирование

Методы резервного копирования

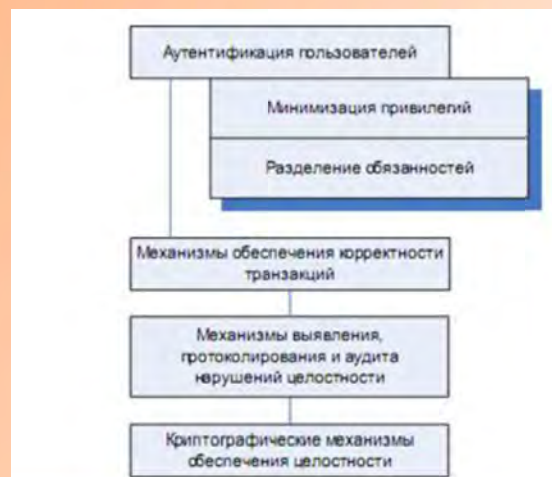
- Полное копирование - все файлы/информация, подлежащая РК переносятся на резервный носитель
- Инкрементальное копирование - РК подвергаются только файлы, изменившиеся с момента предыдущего РК
- Дифференциальное копирование - копируются файлы, изменённые с момента полного резервного копирования.

На практике резервное копирование обычно осуществляется следующим образом: периодически проводится полное резервное копирование, а в промежутках - инкрементальное или дифференциальное. Инкрементальное резервное копирование выполняется быстрее, но дифференциальное РК позволяет легче восстановить оригинал данных.

Алгоритм работы IDS



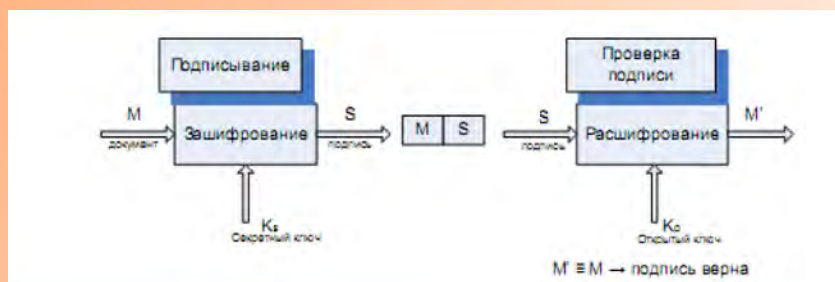
Структура системы защиты от угроз нарушения целостности



Требования к ЭЦП

1. Цифровая подпись должна позволять доказать, что именно законный автор, и никто другой, сознательно подписал документ.
2. Цифровая подпись должна представлять собой неотъемлемую часть документа. Должно быть невозможно отделить подпись от документа и использовать её для подписывания других документов.
3. Цифровая подпись должна обеспечивать невозможность изменения подписанного документа (в том числе и для самого автора!).
4. Факт подписывания документа должен быть юридически доказуемым.
5. Должен быть невозможным отказ от авторства подписанного документа.

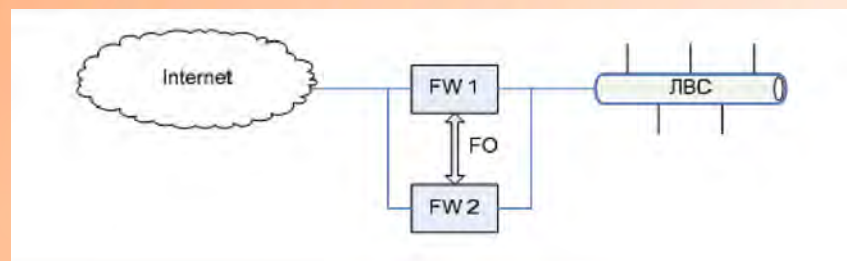
Механизм защиты ЭЦП



Структура системы защиты от угроз нарушения доступности



Дублирование МЭ (вариант реализации)



Методы резервного копирования

1. **Полное** /full/. В этом случае все без исключения файлы, потенциально подвергаемые резервному копированию, переносятся на резервный носитель.
2. **Инкрементальное** /incremental/. Резервному копированию подвергаются только файлы, изменённые с момента последнего инкрементального копирования.
3. **Дифференциальное** /differential/. Копируются файлы, изменённые с момента полного резервного копирования. Количество копируемых данных в этом случае с каждым разом возрастает.

Типы RAID-массивов

Уровень 0. В данном случае несколько дисков представляются как один виртуальный диск. Защита от сбоев на данном уровне никак не обеспечивается.

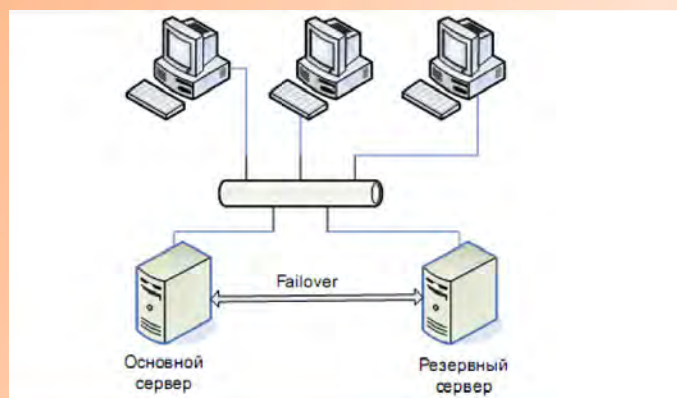
- Уровень 1. Реализуется зеркалирование – идентичные данные хранятся на нескольких (обычно на двух) дисках. Данный вариант обеспечивает надёжную защиту от сбоев носителя, однако является чрезвычайно неэффективным.

- Уровень 2. Биты данных поочерёдно размещаются на различных дисках; имеются выделенные диски, содержащие контрольные суммы. Для контроля ошибок используется код Хэмминга. Используется крайне редко.

Типы RAID массивов

Уровни 3,4 Байты или блоки данных записываются на различные диски, биты чётности – на выделенный диск.
- Уровень 5 Данные и контрольные суммы распределяются по всем дискам. Возможно одновременное выполнение нескольких операций чтения или записи, что значительно повышает общую производительность системы.
- Уровень 7 Функционирование аналогично массивам уровня 5, дополнительно на аппаратном уровне реализовано представление массива в виде единого виртуального диска.

Дублирование серверов



След занятие - зачёт, будут высланы Ефимову