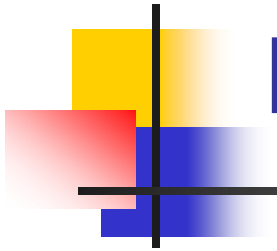
Метод генетической оптимизации схем ролевого доступа к информации

Саенко И.Б. и **Котенко И.В.**

Учреждение Российской академии наук
Санкт-Петербургский институт
информатики и автоматизации РАН
(СПИИРАН)



План (1)

- ✓ **Введение**
- Формальная постановка задачи
- Реализация ГА
- Оценка
- Заключение

Понятие проблемы RMP

[R. S. Sandhu , etc., 1996] ввел модель ролевого управления доступом (Role-Based Access Control – **RBAC)**, чтобы упростить централизованное управление доступом к информации при большом количестве пользователей.

[E. J. Coyne , 1995]: практическое применение **RBAC** привело к необходимости решения проблемы, названной ***инжинирингом ролей*** (*role engineering* - **RE). RE** – это задача определения полного и корректного множества ролей и связанных с ним полномочий.

[J. Vaidya , etc., 2006]: существуют 2 основных подхода к выполнению **RE**: ***нисходящий*** (*top-down*) и ***восходящий*** (*bottom-up*).

[M. Kuhlmann , etc., 2003]: восходящий подход назван ***извлечением ролей*** (*role mining*), так как основными элементами в нем являются методы ИАД (*data mining*). Процесс *Role Mining* является слишком сложной проблемой, которая названа ***проблемой извлечения ролей*** (*role mining problem* - **RMP).**

Проблема RMP является NP-полной!!!

Обзор работ (1)

[D. Zhang , etc., 2007] предложил ***графовые*** подходы для решения проблемы RMP. Однако графовые модели трудны для применения в многокритериальных случаях.

[A. Colantonio, etc., 2008] предложил ***стоимостной*** подход к решения проблемы RMP, но он, обладает существенными ограничениями.

[M. Frank , etc., 2008] предложил ***вероятностный*** подход к проблеме RMP, но он, имеет низкую точность и надежность.

Обзор работ (2)

[J. Vaidya, etc., 2007] предложил ***кластерный*** подход к решению проблемы RMP, но не верится, что он имеет широкую область применения

[H. Roeckle, etc., 2000; M. Strembeck, etc., 2002; I. Molloy, etc., 2008] предложил ***семантико-ориентированный*** подход к решению RMP, но он треьует дополнительного знания семантики.

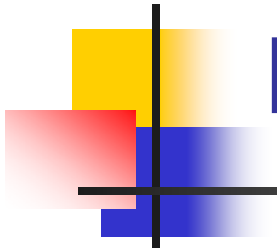
[H. Lu, etc., 2008] предложил решать **RMP** как задачу оптимальной ***декомпозиции булевых матриц***, но он не приемлем для многих случаев.

Обзор работ (3)

[N. Hu, etc., 2006] применил генетические алгоритмы (**ГА**) для решения задач информационной безопасности, но ГА использованы ***совместно с*** RBAC, а не для ***создания*** RBAC.

[N. Semmanche, etc., 2000] успешно применил ГА для управления доступом в ***сервисах Web***, но эти сервисы не используют RBAC .

Использование ГА для решения RMP также возможно !!!



План (2)

- Введение
- **Формальная постановка задачи**
- Реализация ГА
- Оценка
- Заключение

Постановка задачи

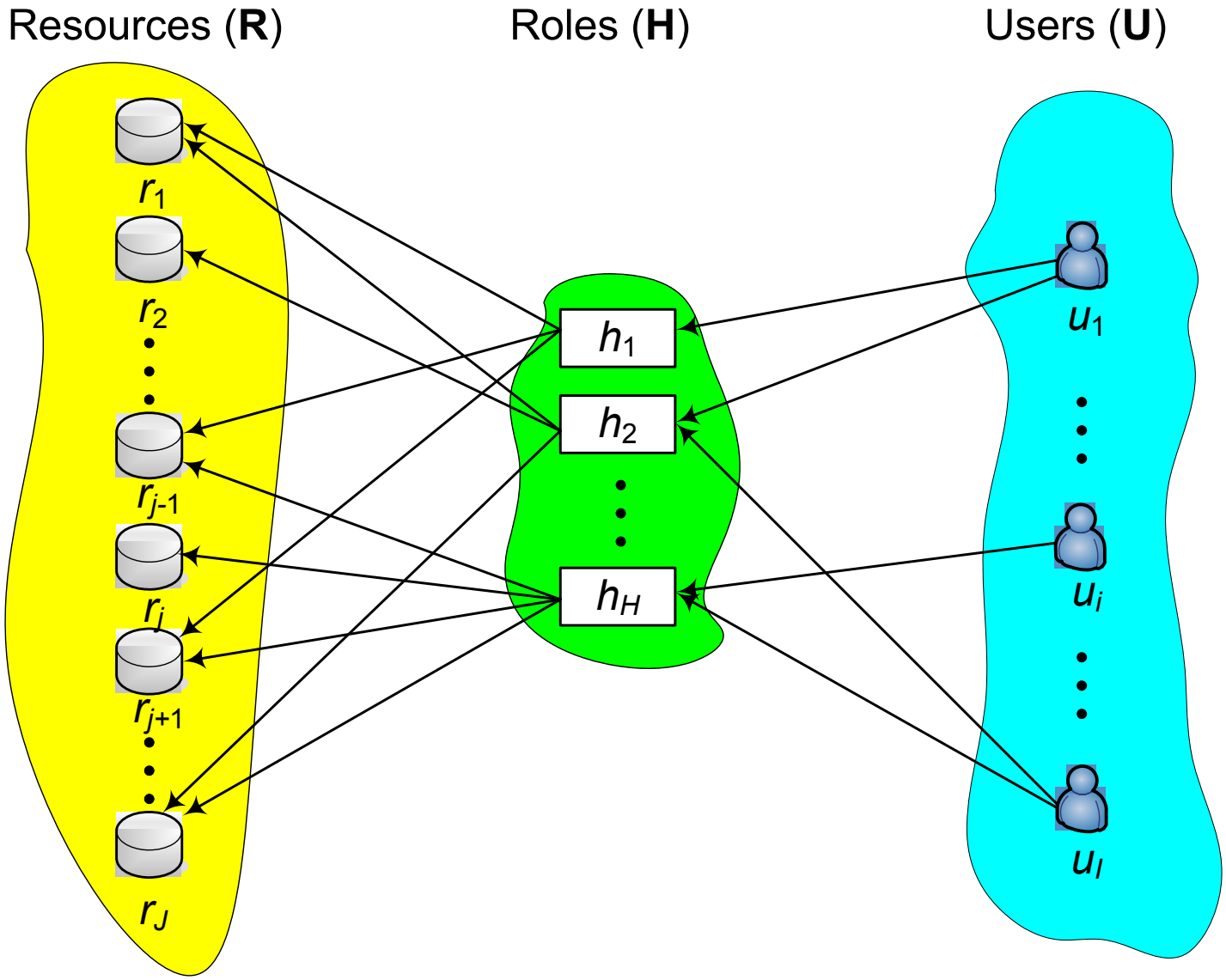
Дано:

- множество *пользователей*;
- множество системных *ресурсов*;
- а *“требуемая” матрица,* определяющая желаемые полномочия по доступу пользователей к ресурсам

Требуется найти:

- минимально необходимое *количество ролей*;
- матрицу *“пользователи-роли”;*
- матрицу *“роли-ресурсы”*

Представление RBAC



Исходные данные

Множество *пользователей*	$\mathbf{U} = \{u_i\}$
Множество *ресурсов*	$\mathbf{R} = \{r_j\}$
Множество *возможных ролей*	$\mathbf{H} = \{h_k\}$
“Требуемая” матрица контроля доступа	$\mathbf{S}^{\text{req}} = \{s_{ij}^{\text{req}}\} \subseteq \mathbf{U} \times \mathbf{R}$

Переменные

Матрица *роли-ресурсы*	$\mathbf{X} = \left\{ x_{kj} \right\}, j = \overline{1, J}$ $k = \overline{1, K}$
Матрица *пользователи-роли*	$\mathbf{Y} = \left\{ y_{ik} \right\}, i = \overline{1, I}$ $k = \overline{1, K}$

Целевые функции

"Реальная" схема контроля доступа: $\mathbf{S}^{\text{real}} = \left\{ s_{ij}^{\text{real}} \right\}$

$$\mathbf{S}^{\text{real}} = \mathbf{Y} \cdot \mathbf{X} \quad (1) \qquad s_{ij}^{\text{real}} = \sum_{k=1}^{K} y_{ik} x_{kj} \quad (2)$$

Первая целевая функция G^{conf}

$$G^{\text{conf}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \max\left(0, \left(s_{ij}^{\text{real}} - s_{ij}^{\text{req}} \right)\right) \quad (3)$$

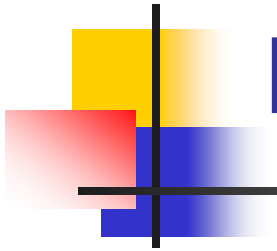
Вторая целевая function G^{accs}

$$G^{\text{accs}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \max\left(0, \left(s_{ij}^{\text{req}} - s_{ij}^{\text{real}} \right)\right) \quad (4)$$

Обобщенный критерий задачи:

$$K \Rightarrow \min;\ G^{\text{conf}} = 0;\ G^{\text{accs}} = 0 \quad (5) \qquad \text{// "базовая" RMP}$$

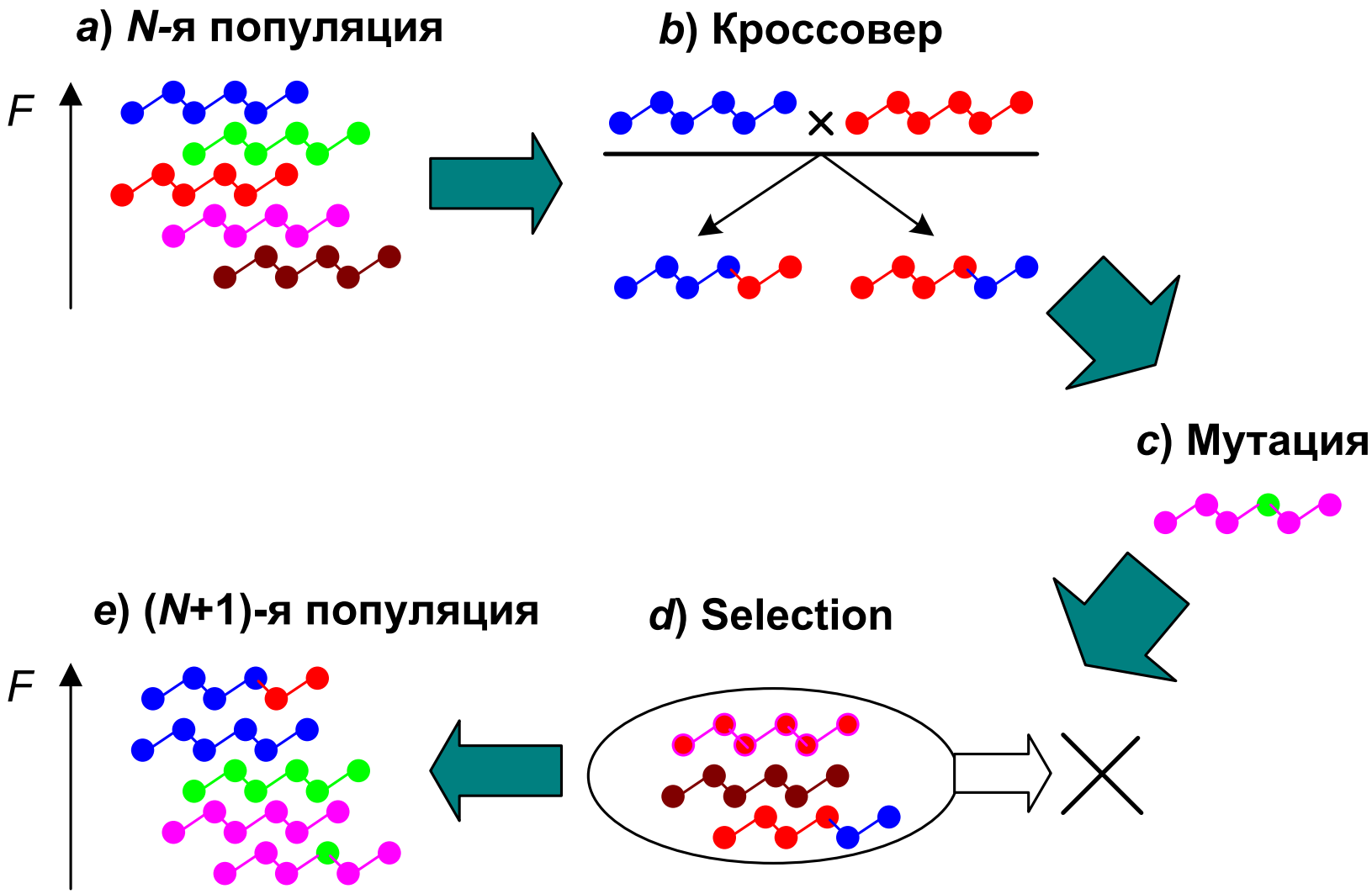
$$\left(|\mathbf{X}| + |\mathbf{Y}| \right) \Rightarrow \min;\ G^{\text{conf}} = 0;\ G^{\text{accs}} = 0 \quad (6) \qquad \text{// "краевая" RMP}$$



План (3)

- Введение
- Формальная постановка задачи
- **Реализация ГА**
- Оценка
- Заключение

Что такое ГА?


a) N-я популяция
b) Кроссовер
c) Мутация
d) Selection
e) (N+1)-я популяция
F
F

Хромосомы (кодирование переменных) (1)

1-я хромосома состоит из столбцов матрицы ***Роли-Ресурсы*:**

Матрица **X**:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & & 0 \\ 0 & 1 & & 0 \\ .. & .. & & .. \\ 1 & 0 & & 1 \\ 0 & 0 & \cdots & 1 \\ 1 & 0 & & 0 \\ .. & .. & & .. \\ 1 & 1 & & 1 \end{bmatrix}$$

Хромосома **X**:

$$Chr(\mathbf{X}) = < \begin{bmatrix} 1 \\ 0 \\ .. \\ 1 \\ 0 \\ 1 \\ .. \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ .. \\ 0 \\ 0 \\ 0 \\ .. \\ 1 \end{bmatrix}, \cdots, \begin{bmatrix} 0 \\ 0 \\ .. \\ 1 \\ 1 \\ 0 \\ .. \\ 1 \end{bmatrix} >$$

Хромосомы (2)

2-я хромосома состоит из столбцов матрицы ***Пользователи -Роли*:**

Матрица **Y**:

$$\mathbf{Y} = \begin{bmatrix} 1 & 1 & & 0 \\ .. & .. & & .. \\ 0 & 0 & & 1 \\ .. & .. & ... & .. \\ 0 & 1 & & 1 \end{bmatrix}$$

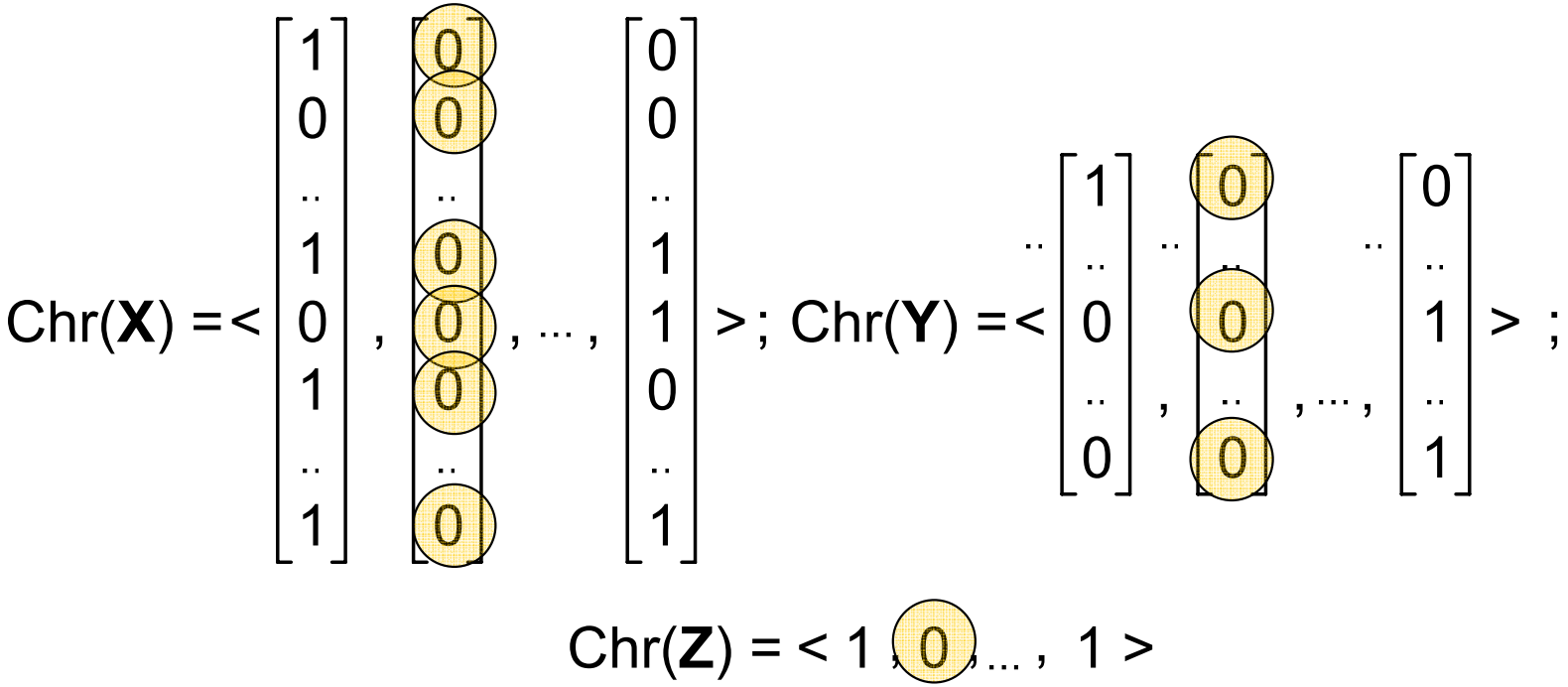
Хромосома **Y**:

$$Chr(\mathbf{Y}) = < \begin{bmatrix} 1 \\ .. \\ 0 \\ .. \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .. \\ 0 \\ .. \\ 1 \end{bmatrix}, ..., \begin{bmatrix} 0 \\ .. \\ 1 \\ .. \\ 1 \end{bmatrix} >$$

Хромосомы (3)

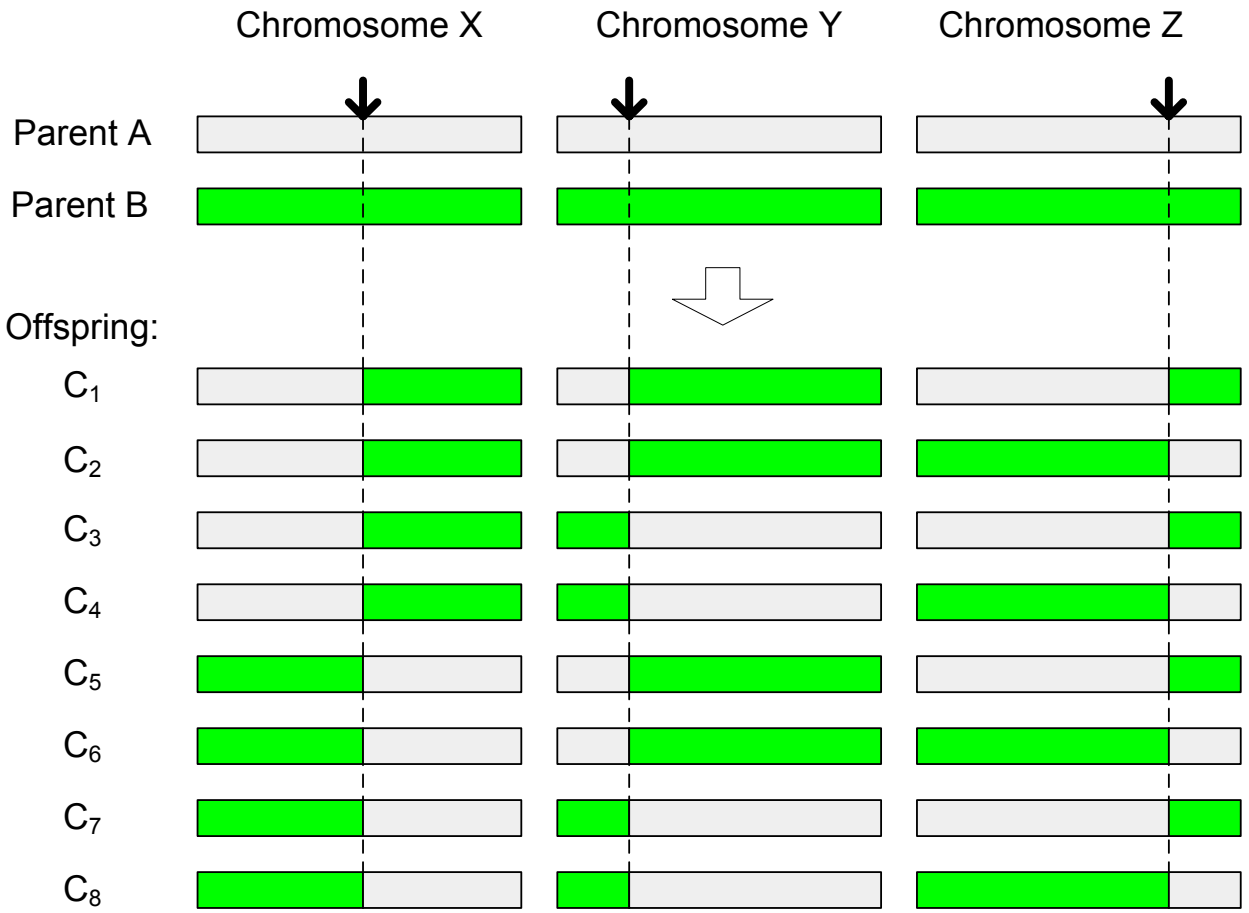
3-я хромосома Chr[**Z**] является *управляющей* и является *бинарной строкой*.

Если ген этой хромосомы равен 1, то соответствующая роль является *активной*, иначе – *пассивной*.



Пример: Роль *r*2 является *пассивной*

Кроссовер



Примечание: новые дочерние хромосомы Chr(**X**) и Chr(**Y**) будут корректировать свои гены в соответствии с информацией, содержащейся в Chr(**Z**).

Мутация (1)

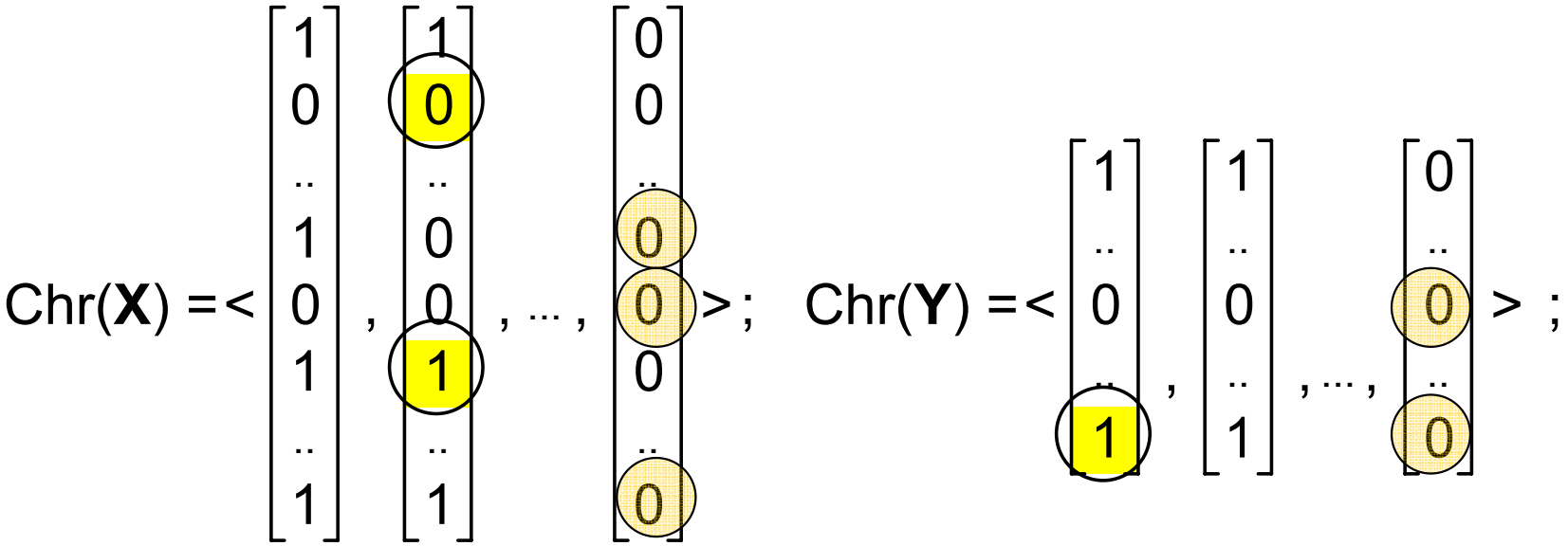
До мутации (выбранные элементы - x_{22}, $x_{2,j+1}$, and $y_{1,l}$):

$$\text{Chr}(\mathbf{X}) = < \begin{bmatrix} 1 \\ 0 \\ .. \\ 1 \\ 0 \\ 1 \\ .. \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ \mathbf{1} \\ .. \\ 0 \\ 0 \\ \mathbf{0} \\ .. \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ .. \\ 1 \\ 1 \\ 0 \\ .. \\ 1 \end{bmatrix} >; \quad \text{Chr}(\mathbf{Y}) = < \begin{bmatrix} 1 \\ .. \\ 0 \\ .. \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} 1 \\ .. \\ 0 \\ .. \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ .. \\ 1 \\ .. \\ 1 \end{bmatrix} > ;$$

$$\text{Chr}(\mathbf{Z}) = < 1 , 1 , \dots , \mathbf{1} >$$

Мутация (2)

После мутации (выбранные элементы - x_{22}, $x_{2,j+1}$, and $y_{1,l}$):



Chr(**Z**) = < 1 , 1 , ... , 0 >

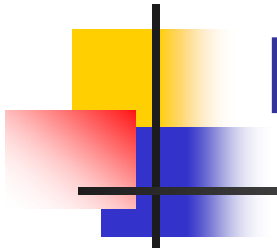
Функция пригодности

“Базовая” RMP:

$$F_{\text{basic}} = \left(\alpha_1 \cdot K + \alpha_2 \cdot G^{\text{conf}} + \alpha_3 \cdot G^{\text{accs}} \right)^{-1} \qquad (7)$$

“Краевая” RMP:

$$F_{\text{edge}} = \left(\alpha_1 \cdot \left(|\mathbf{X}| + |\mathbf{Y}| \right) + \alpha_2 \cdot G^{\text{conf}} + \alpha_3 \cdot G^{\text{accs}} \right)^{-1} \qquad (8)$$

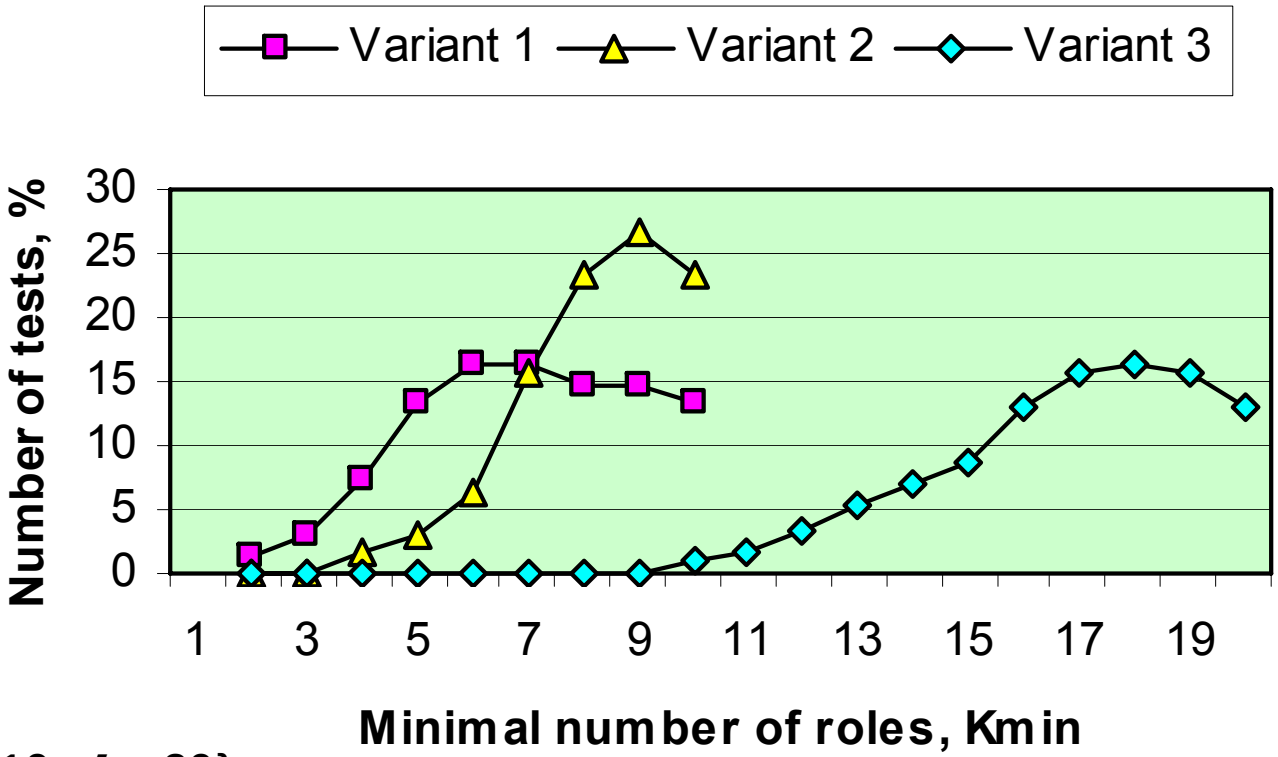


План (4)

- Введение
- Формальная постановка задачи
- Реализация ГА
- **Оценка**
- Заключение

Минимальное количество ролей

Зависимость показывает, как количество *успешных тестов* (в процентах), которые приводят к результату K_{min} как решению задачи, *зависит от* K_{min}.

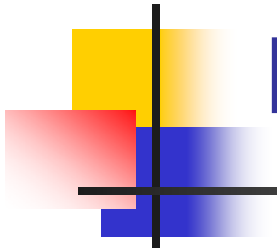


Вариант 1: { *I* = 10, *J* = 20},
Вариант 2: { *I* = 10, *J* = 50},
Вариант 3: { *I* = 20, *J* = 50}.
Вид функции пригодности - ***"базовая"*** RMP.

Оценка производительности

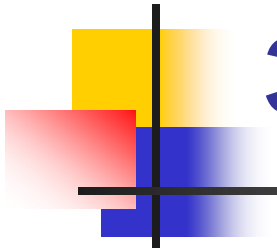
Варианты	*I*	*J*	“Базовая" RMP		“Краевая" RMP	
			$N_{\min}^{pop}$	$N_{\max}^{pop}$	$N_{\min}^{pop}$	$N_{\max}^{pop}$
Вариант 1	10	20	15	30	17	29
Вариант 2	10	50	21	52	20	55
Вариант 3	20	50	48	95	52	105

$N_{\min}^{pop}$ и $N_{\max}^{pop}$ есть *минимальные* и *максимальные* значения N^{pop}, которые были получены в тестах.



План (5)

- Введение
- Формальная постановка задачи
- Реализация ГА
- Оценка
- **Заключение**



Заключение

- Использование ГА в *стандартной* форме связано со значительными трудностями. Основная причина этого заключается в *переменной длине* хромосом.
- Чтобы преодолеть эту проблему, мы предлагаем несколько *оригинальных решений*, касающихся реализации ГА (см. выше).

- Будущие работы связаны с реализацией ГА для решения RMP для случая *иерархии ролей*.

Контактная информация

Саенко Игорь Борисович (СПИИРАН)
saenko@comsec.spb.ru
http://comsec.spb.ru/saenko/
Котенко Игорь Витальевич (СПИИРАН)
ivkote@comsec.spb.ru
http://comsec.spb.ru/kotenko/

Благодарности

- Работа выполняется при финансовой поддержке РФФИ (проекты №10-01-00826 и 11-07-00435-а), программы фундаментальных исследований ОНИТ РАН (проект № 3.2) и при частичной финансовой поддержке, осуществляемой в рамках проекта Евросоюза Massif.