

## ИНТЕЛИГЕНТНА КОМПЮТЪРНА СИСТЕМА ЗА НЕПРЕКЪСНАТА ОЦЕНКА И ПОДДЪРЖАНЕ НА УПРАВЛЕНИЕТО НА РИСКОВЕТЕ ЗА БЕЗОПАСНОСТ В МИННИ ПРЕДПРИЯТИЯ

Тодор Петров

Минно-геоложки университет  
"Св. Иван Рилски"  
София 1700, България  
E-mail: [tpp@mgu.bg](mailto:tpp@mgu.bg)

### РЕЗЮМЕ

В материала се представя теоретичния фундамент и структура на интелигентна система за непрекъсната оценка на рисковете и поддържане на управлението на безопасността в минни предприятия. Системата е базирана на доверителна мрежа на Бейз конструирана за извличане на изводи за нивото на безопасност. За оценка на варианти на решения за подобряване на безопасността предложената мрежа на Бейз е развита до диаграма на въздействията. Дискутира се връзката на системата с външни бази данни, резултати от тестване на системата с реални данни, нейната работоспособност и възможности за обучение в условията на неопределеност.

### ВЪВЕДЕНИЕ

Информационното осигуряване на управлението на безопасността изисква обработка на голям обем данни от качествен и количествен характер. Практиката е показала, че отчитането само на количествените показатели, като честота на инцидентите, коефициенти и индекси на честота и тежест и средните оценки не са достатъчни за характеризирани на нивото на безопасност в предприятия и техни подразделения и обекти. За борба с опасностите е необходима и качествена система за оценка на рисковете. За важността на проблема красноречиво говори инициативата на ISO да разработи нова серия родови стандарти за качествено управление на безопасността ISO 18000. Днес, разследването и регистрирането само на един инцидент се документира в минимум в 60 полета с данни от различен формат. Изследването на рисковете в големите минни предприятия нормално обхваща над 3000 масива от данни за описание на 50 и повече инцидента на година. и преминава през класифицирането на всеки инцидент в 21 (таб. 1) показателя всеки от които се определя с 2 до 26 групи характеристики (*Михайлов и Петров, (1997), Михайлов (2001)*). От психологията и когнитивната наука е известно, че човешкото съзнание не е пригодно да обработва такива структури а подхожда чрез абстракция и йерархично подреждане на наличната информация, като среща сериозни затруднения при обхващането на повече от три алтернативи при вземане на решение. Подобно информационно претоварване на съзнанието води до включване на механизма на игнориране на информация и евристично отсичане на варианти, което значително повишава риска от неправилни решения и може да доведе до опасни действия с тежки последици. Проблемът е особено актуален при критични по време ситуации. Друга съществена характеристика на

дейността по оценка и управление на рисковете за безопасност, е че решенията се взимат в условия на вероятностна и размита неопределеност, което на практика означава, че очакваните ефекти от приложените мероприятия също ще има вероятностен характер. Така също отчитането на експертни мнения е ежедневие в дейността на специалистите по безопасност и ефективното имплементиране на експертния опит в управлението е от съществено значение. Типично прилагания подход на изследване на рисковете, чрез самостоятелното изучаване на влиянието на отделни факторни групи е свързан неминуемо със загуба на ценна информация за взаимовръзките в системата за безопасност. Такова "разглобяване" представлява необратим процес от информационна гледна точка.

В светлината на казаното изпъква необходимостта от прилагането на синергетичен подход при изследването на рисковете за безопасност, който да държи сметка за спецификите на изследвания обект. Реализацията на подхода във вид удобен за практическо използване трябва да стъпи върху интелигентна компютърна система. Интелигентните компютърни системи работят с бази знания. Успешно формализиране на база знания за безопасност може да се осъществи само за специфична предметна област - домейн. Практиката е показала, че всички успешни интелигентни системи, са разработени за отделни домейни. Пример за такива са експертните системи за медицинска диагностика още в зората на изкуствения интелект през 70-те години.

За реализиране на този подход е разработена интелигентна компютърна система MAR.NET (My Accident Risk dot Net).

## ПОКАЗАТЕЛИ ХАРАКТЕРИЗИРАЩИ РИСКОВЕТЕ ЗА БЕЗОПАСНОСТ ИЗПОЛЗВАНИ В MAR.NET

В най-добрата световна практика по безопасност съществува стремеж към създаване на единна система от показатели характеризиращи рисковете. Такава система е предложена за минните предприятия в *Михайлов и колектив (2002)* и е представена в таблица 1.

Таблица 1. Показатели характеризиращи рисковете.

Наименование	Етикет
Период в денонощието (час)	01.Hour
Групи работници – професия	02.Occupation
Степен на образование	03.Education
Общ трудов стаж	04.Practice
Трудов стаж в предприятието	05.Practice_Co
Трудов стаж по посочената професия	06.Practice_Pro
Пореден работен ден след последната почивка	07.Day_after
Време след започване на работата	08.Hour_after
Място на злополуката	09.Place
Вид на извършваната работа от пострадалия	10.Job
Вид на произшествието довело до злополуката	11.Incident
Човешки фактор в причините за злополуката	12.Human_Factor
Материален фактор в причините за злополуката	13.Material_Factor
Опасности на околната среда оказали влияние	14.Environment
Отклонение от нормалните действия и условия	15.1.Deviation_A 15.2.Deviation_E
Тежест на злополуката	16.Severity
Увредени части и системи на тялото	17.Body
Вид на увреждането	18.Injure
Период на възстановяване	19.Recover_Period
Мерки за редуциране на риска	20.Measure
Машини имащи отношение към злополуката	21.Machinery

Характеристиките на предложените показатели, не са представени в материала поради липса на място.

Така дефинирана системата от показатели може да се използва за качествено изследване на рисковете за безопасност и във всички други браншове. За изследването на рисковете обаче, разумния компромис е да се строят единни системи от показатели и характеристики на ниво бранш. Конвергенцията на системите на глобално ниво е следваща стъпка.

## ИЗВЛИЧАНЕ НА ИЗВОДИ ЗА НИВОТО НА БЕЗОПАСНОСТ

Извличането на заключения за нивото на безопасността в даден обект, предприятие или бранш е на практика процес на формиране на статистически извод за синергетичното въздействие на отчетените фактори върху тежестта на инцидентите в условия на неопределеност.

Нивото на безопасност е пряко свързано с определянето на риска от инциденти. Класическата дефиниция на риска в безопасността е произведението:

$$\text{РИСК} = \text{ВЕРОЯТНОСТ} \times \text{ПОСЛЕДИЦИ}$$

При описанието на рисковете за безопасност в даден обект трябва да влязат всички възможни рискове. След изпълнение на програмите по безопасност обект на оценка е текущия - остатъчен риск  $R_c$ . Когато се говори за ниво на безопасност обикновено се има в предвид база за сравнение на текущо ниво с допустимо ниво – социално хуманно, икономическо. Като количествена връзка между социално хуманните и икономическите последици може да се посочи броят на изгубените човеко дни. Показателя 19.Recover\_Period характеризира рисковете за безопасност чрез дискретно вероятностно разпределение инцидентите да причинят загуби попадащи в 10 интервални множества загубени човекодни – “А. До 3 дни”, “В. 4 до 17 дни”, ... , “J. 6000 дни (невъзстановими злополуки)”.

Нивото на безопасността следва да служи, като количествена мярка на текущите рискове.

Ако се оцени минималния прагов риск, под който системата за безопасност не прави отчети при вероятност  $1/10^8$  и последици – до 3 дни на  $R_0 = 3 \times 10^{-8}$  то нивото на безопасност може да се дефинира със зависимостта (1).

$$L_s = \log(R_c / R_0) \quad (1)$$

Тя притежава някои свойства, които я правят удобна за оценяване на ниво. Първо, тъй като рисковете са винаги положителни величини  $R_c \geq R_0 > 0$ , стойността на логаритъма винаги ще може да бъде изчислена и ще бъде положителна величина  $L_s > 0$ . Второ, когато текущият риск стане равен на праговият  $R_c = R_0$ , нивото  $L_s = 0$ . И не на последно място, от психофизическа гледна точка възприятата за ниво са свързани именно с логаритмична зависимост обобщена от закона на Вебер и Фехнер. От тези съображения се препоръчва логаритъма в (1) да бъде десетичен.

Освен от информацията за нивото на безопасността, управлението се нуждае и от детайлно качествено изследване на рисковете. Натрупването на знания за работата на системата за осигуряване на безопасност се базира на изследването на извадкова информация (регистър на трудовите злополуки, книга за аварияте, протоколи от инспекции и др.). Това води неумолимо до необходимостта от статистическа оценка на изводите за реалната система, която очевидно е по богата на свойства от наблюдаемата (*Hines, (2000); Вентцель (2001)*). Нелинейния характер на динамиката на опасностите и възможността системата да изпадне в хаотичен режим (*Guastello (1997); Stengers, Prigogine (1997); Петров (1999)*) поставят проблема извън областта на приложение на добре разработени методи за намаляване на размерността на задачата, като детерминистичния

факторен анализ и класическите статистически средни оценки (Трухаев, Горшков (1985)).

### Подход на Бейз за извличане на извод

Риска в безопасността се дефинира като произведение от вероятност по последици. Честотната интерпретация на вероятността често се нарича *обективна* или *класическа гледна точка*. В статистическата теория подходът на Бейз за извличане на извод предлага различна интерпретация на вероятността, наречена *субективна гледна точка*. Основно място в подхода заема идеята за условната вероятност. Подходът на Бейз изисква от изследователя да използва субективна вероятност, като мярка за неговата степен на увереност за състоянието на наблюдавания обект. Степента на увереност се специфицира с вероятностно разпределение за даден непознат параметър. Тази процедура е напълно различна от всички други статистически подходи. В другите статистически подходи, непознатите параметри се третираат, като непознати константи. Подходът на Бейз изисква от нас да мислим за непознатите параметри, като за случайни променливи.

### Закон за пълната вероятност и теорема на Бейз

Оценката на съвместното влияние на събитията довели до инцидент е в основата на детайлното изследване на рисковете. Основните зависимости на които е построен модела MAR.NET могат да се дефинират както следва. Нека  $A_1, A_2, \dots$  е изброима колекция от събития, която поделва едно и също пространство на реализациите –  $S$  така, че събитията са взаимно независими и тяхното обединение е  $S$ . Нека  $B$  е друго събитие и да допуснем, че ни е известно  $P(A_i)$  и  $P(B|A_i)$  за  $i \in I$ .

Тогава може да се покаже, че:

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots, \quad (2)$$

резултат известен като закон за пълната вероятност;

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots}, \quad (3)$$

резултат известен като теорема на Бейз;

$$P(A_1, A_2, \dots, A_n) = P(A_1|A_2, A_3, \dots, A_n)P(A_2|A_3, A_4, \dots, A_n) \dots P(A_{n-1}|A_n)P(A_n), \quad (4)$$

резултат известен, като верижно правило, особено важно при мрежите на Бейз.

### Бейзови доверителни мрежи - БДМ

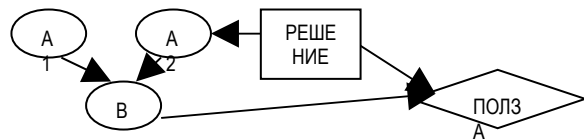
Мрежата на Бейз се представя с насочен ацикличесен граф със следните свойства: възлите представят случайни променливи и се наричат още шансови, ребрата представят вероятностни зависимости между променливите. На тези зависимости се дава количествен израз чрез множество от условни вероятностни таблици (увт). За променливи без родители разпределението е безусловно наричано още периферно. Основното предимство пред "класическите статистически заключения" (които работят по-скоро с доверителни интервали, от колкото с изявления на вероятности) е, че Бейзовото заключение пълно описва

фактът, че очакването само не може да предскаже вероятността от неочаквани събития, без предварително съществуваща информация за последните. Необходимостта да се приеме приорно мнение е ключова част от Бейзовото заключение. Разбира се това изискване е слабост. Не винаги е лесно да се получи приорна вероятност, освен от експерти.

**Извличане на заключение в мрежите на Бейз** Изводът или заключението в мрежите на Бейз означава изчисляването на условната вероятност за някой променливи да е дадена информация (указание) от други променливи. Това е лесно, когато всички указания са от променливи, предшественици на променливите, които ни интересуват ( $B_i \rightarrow A$ , фиг. 1) Но когато указание се даде от потомък към променлива от интерес, ние трябва да извлечем извод срещу посоките на ребрата. За целта използваме теоремата на Бейз (3).

### Диаграми на въздействията и вземане на решения

Модел за поддържане вземането на решения може да се конструира и на базата на чиста БДМ, но концепциите за функция на ползата и решенията не са ясно формулирани и напълно покрити. Разширението на БДМ с два типа възли – представящи решенията и ползите представлява диаграма на въздействията (фиг. 2). Диаграмите на въздействията се използват за оценка на различни варианти на решения чрез изчисляване на очакваните ползи от предприетите действия.

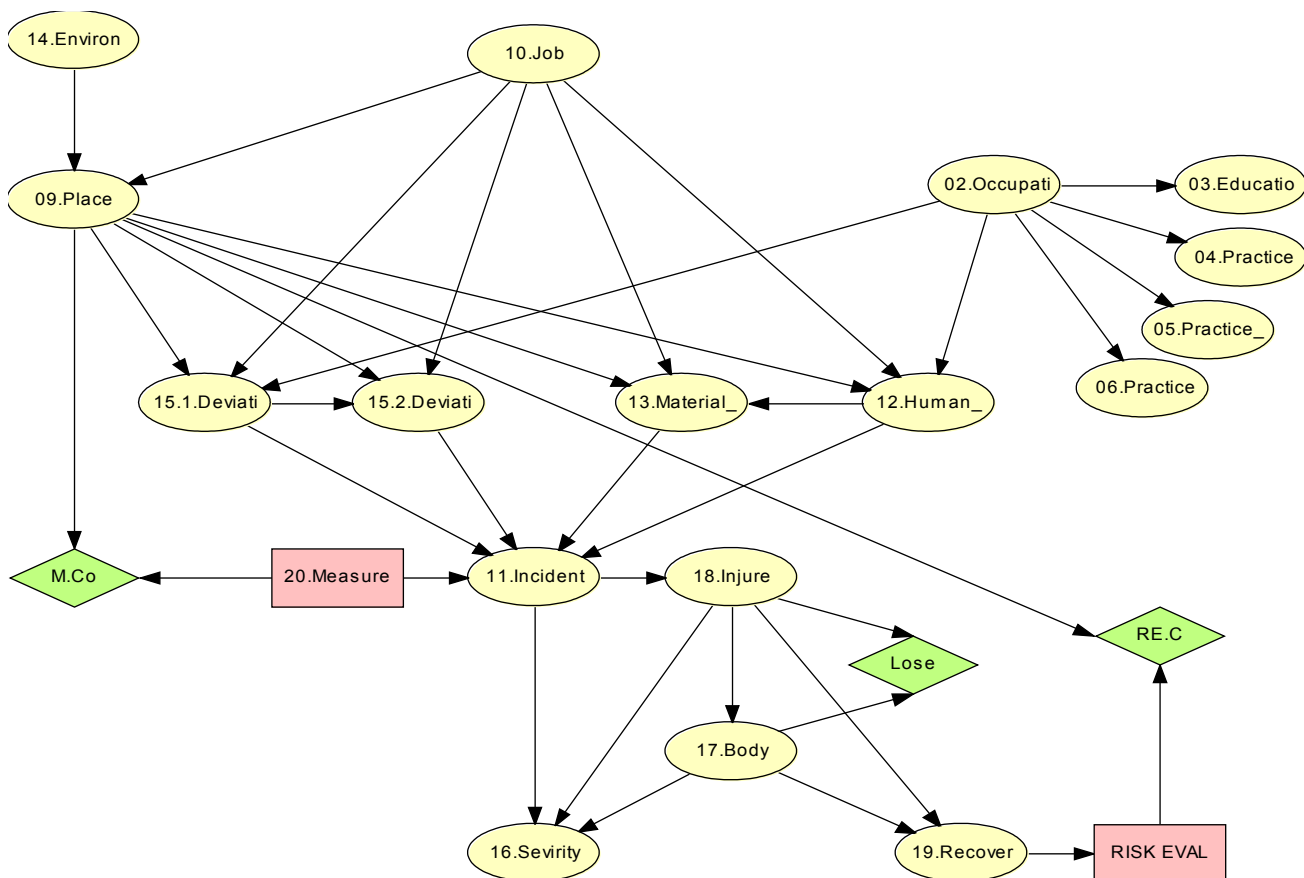


Фигура 1. Примерна мрежа на развита до диаграма на въздействията с възел-решение и възел-полза

Решенията в мрежата трябва да са свързани верижно в логична последователност. Възлите от които зависят решенията, трябва да са с известни състояния преди решението да бъде взето.

### СТРУКТУРА НА MAR.NET

MAR.NET е изградена като доверителна мрежа на Бейз развита до диаграма на въздействията. Нейната структура е дадена на фиг. 2. Основното предназначение на MAR.NET системата е поддържане на решения за повишаване на нивото на безопасността в индустриален обект, поделение, предприятие и цял бранш. Състоянията на възлите са адаптирани за спецификата на минния бранш. Мрежата се състои от шансови възли, които могат да приемат състояния, така както характеристиките на показателите дадени в табл. 1. Условната зависимост между променливите може директно да се прочете от графа на фиг. 2. Вероятностите на отделните състояния на възлите в мрежата определят бейзовата увереност да се случи инцидент при зададените от родителските възли условия.



Фигура 2. Модела - MAR.NET за оценка на решения за подобряване на безопасността реализиран на базата на бейзова доверителна мрежа развита до диаграма на въздействията

Икономическия риск от последиците се оценява от възела "Loses". Рискът измерен в изгубени работни дни се оценява от възела 19.Recover\_Period. Социално хуманната тежест на инцидента се оценява от възела 16.Severity. И трите възела са наследници на родителски възли и свързаните с тях вероятностни разпределения са условни.

Периферните възли в корена на мрежата са 14.Environment, 10.Job и 02.Occupation. При инициализиране на мрежата вероятностите на състоянията се разпределят униформено. Нулевите вероятности присвоени на състояния на възли означават изключване на възможността от изява на тези състояния. Вида на вероятностните таблици на периферните възли са от вида на табл. 2.

Таблица 2. Вероятностна таблица на възел 10.Job.

Състояние	Вероятност
А. Транспортни и товарни работи	0.2
В. Нормална експлоатация	0.2
...	...
Е. Други	0.2

Влиянието на възлите 01.Howeg-"Период в денонощието", 07.Day\_after-"Ден след почивка" и, 08.Howeg\_after-"Време след започване на работа" (таб. 1) са обект на допълнително изследване и на този етап не са свързани в мрежата. Останалите възли в мрежата се

характеризират с условни вероятностни разпределения (таб. 3). Така например възел 17.Body е наследник на възел 18.Injure и се явява родител на още два шансови възела: 16.Severity, 19.Recovery\_Period и на възела на загубите – Loses.

Състоянията на променливите на възлите образуват независима група с пълна вероятност 1.

Таблица 3. Условна вероятностна таблица P(17.Body| 18.Injure) инициализирана с униформено разпределение

18.Injure	A	B	...	Z
17.Body				
А. Глава	0.25	0.25	0.25	0.25
В. Ръце	0.25	0.25	0.25	0.25
С. Крака	0.25	0.25	0.25	0.25
Д. Други	0.25	0.25	0.25	0.25
Сума	1	1	1	1

В сърцето на мрежата е шансовия възел 11.Incident – "Вид на произшествието довело до злополуката". В реализацията на мрежата, възелът е пряк наследник на пет родителя - четири шансови възела и един-възел решение съответно 12,13, 15.1, 15.2 и 20 (таб. 1). Свързаната с него условна вероятностна таблица предстатява многомерен масив зададен от условното вероятностно разпределение:

$P(11.Incident|12.Human\_Factor,13.Material\_Factor,$   
 $15.1.Deviation\_Environment,15.2.Deviatin\_Action,(5)$   
 $20.Measure)$

Размерността на масива е  $18*14*15*5*2 = 37800$ .

Възелът "Loses" очакваните загуби  $P(Loses|Injure,Body)$  при реализация на различни конфигурации от състояния на променливите "Вид на увреждането" и "Увредени части" на тялото (табл. 4). За по-голяма прецизност, в условното разпределение може да се добави и периода на възстановяване -  $P(Loses|18.Injure,17.Body,19.Recovery\_Period)$ .

Таблица 4. Таблица на ползите на възела "Loses"

Body	A			B ...	
Injure	A	...	P	A	...
Loses	-90	...	-100000	-150	

Постоянната оценка на ефекта от изпълнението на програмите по безопасност се извършва чрез свързването на възела решение – 20. Measure – "Мерки за редуциране на риска". За конкретния пример е прието възелът да има само две състояния: Action 0 и Safety Program. Това отговаря на вземане на решение да не се предприемат действия или да се приведе в изпълнение програма по безопасност. Очаквания ефект от решенията ще се отчете чрез разпространение на вероятностите върху състоянията на възела 11.Incident. Изпълнението на решенията е свързано с разходи. Тези разходи са отразени във възела на ползите "M.Cost" (таб. 5), който е наследник на възлите "09.Place" – Място на инцидента и 20.Measure.

Таблица 5. Таблица на ползите на възел "M.Cost"

Place	A		B		C
Measure	Action 0	Safety Program	Action 0	Safety Program	...
M.Cost	0	-5000	0	-10000	...

В мрежата е включен и възел-решение "RISK EVALUATION" с чиято помощ да се оценят очакваните ползи от изпълнение на процедури по оценка на риска. В дадения модел той има две състояния: "Да" и "Не". Натиск за стартирането на процедурите по оценка на риска оказва разпределението на изгубените работни дни отразени във възела 19.Recovery\_Period. Нарастването на честотата и тежестта на инцидентите ще измества вероятностното разпределение на периода на възстановяване "надясно" към по-дългите интервали. На най-тежките злополуки отговаря състоянието "6000 загубени човекодни". Това води до необходимостта да се вземе решение за оценка на рисковете. От друга страна оценката на риска винаги е свързана с типове работни места. В предложения модел възелът на решението "RISK EVALUATION" е наследник на шансовия възел 19.Recovery\_Period. Разходите свързани с процедурите на оценката са представени с възел на ползите "RE.Cost", чието вероятностно разпределение е условно на родителските възли "RISK EVALUATION" и "09.Place" (фиг. 2, таб. 6).

Таблица 6. Таблица на ползите на възел "RE.Cost"

Place	A		B		C
Risk Evaluation	ДА	НЕ	ДА	НЕ	...
RE.Cost	-1000	0	-2500	0	...

Изборът на алтернативно решение става на принципа на максималната очаквана полза. Глобалната функция на ползите  $U$  е сума от всички локални пози (5). Локалната полза  $u_i$  от алтернатива на решение зависи от вероятностното разпределение на свързаните с него конфигурации  $c_j$  от състояния на променливи (6). Така например условното вероятностно разпределение на ползите във възела Loses съдържа  $j = 104$  вероятности. Но при изчислението на  $U$  полза се отчитат и локалните ползи извлечени от възлите RE.Cost и M.Cost, т.е.  $i = 3$ .

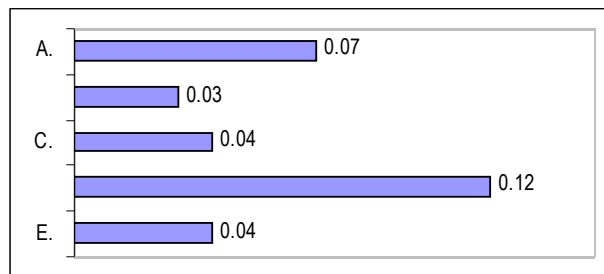
$$U = \sum_i u_i \quad (5)$$

$$u_i = \sum_j p_j u_j \quad (6)$$

където  $p_j$  е условната вероятност за реализиране на конфигурацията  $c_j$ , а  $u_j$  - е стойността на ползата свързана с тази реализация съгласно таблиците на ползите.

#### ИЗВЛИЧАНЕ НА ЗАКЛЮЧЕНИЯ ЗА НИВОТО НА БЕЗОПАСНОСТ С MAR.NET

При получаването на нова приорна информация в мрежата за част от състоянията на възлите следва да се стартира процедурата за извличане на извод. Извършва се разпространение на вероятностите по дъгите на графа като се изчислява постериорното разпределение на цялата мрежа. В резултат може да се наблюдават вероятностите на всяко състояние на възлите и оценката на всяка алтернатива на описаните в мрежата решения в термините на функция на ползите. Един примерен резултат за възела 10.Job-"Вид на извършваната работа" е даден на фиг. 3.



Фигура 3. Постериорно вероятностно разпределение на променливата 10.Job 6).

Обучението на мрежата може да става по 6 различни начина. Първият начин е ръчно въвеждане на вероятностите за отделните състояния на възлите. Отчитайки размерността на дискретните разпределения, това е тежка задача. За целта са разработени специални алгоритми за обучение и адаптиране. Веднага след създаване на мрежата са възможни две начални състояния – всички вероятности да бъдат нула или униформено разпределени за всеки от възлите. От гледна точка на оценката на безопасността по-подходящо е да се допусне начално състояние с униформени разпределения. С други думи не се изключва и не се приоритизира нито едно от възможните състояния. В следващите фази на обучение началните вероятности ще претърпяват промени в съответствие с новополучената информация за спецификата на безопасността в обекта, за който се адаптира MAR.NET. Алтернативен метод за инициализиране на вероятностните разпределения на възлите в MAR.NET е чрез обединение от SQL запитвания към базите данни на MAR. За целта, структурата на мрежата се описва във вида 5 от корените към периферията на графа. SQL запитванията в обединението следват тази последователност. В резултат се формират таблиците с условните вероятностни разпределения на възлите в мрежата.

#### Обучение от данни за регистрирани случаи

Този тип обучение е подходящо за инициализиране на вероятностните разпределения, след дефиниране на структурата на мрежата. Определя се множество от възли, за които се разполага с приорна информация. На всеки възел от множеството се присвоява таблица на опита отчитаща броят на реализациите на всяко състояние. В обучението на MAR.NET участват всички възли (таб. 7)

Таблица 7. Обучаваща съвкупност за MAR.NET

N01	N02	N03	N04	N05	N06	...	N21
A	A	D	Q	A	B	...	N/A
C	I	D	D	N/A	C	...	C
...							

Когато липсва информация за регистрирания случай, съответното поле се бележи с "N/A". Буквените кодове отговарят на имената на състоянията на променливите отразени в мрежата. Удобен начин за извличане на обучаващ пример е да се използват създадени вече бази данни. MAR.NET е разработена съвместимо с базите данни на софтуерният продукт за отчитане и анализ на рисковете за безопасност в мините Mine Accident Risk версия 2002 (MAR). Продуктът се развива от 1995 година в катедра "Руднична вентилация и Техническа Безопасност" при Минно-Геоложки Университет "Св. Иван Рилски".

Методът използван за обучение на MAR.NET е основно използвания при бейзовите мрежи EM-алгоритъм за графично асоциирани модели с липсващи данни (Cowell and Dawid, 1992; Lauritzen, 1995). Целта на алгоритъма е попълването на условните вероятностни таблици свързани с възлите в мрежата. За тази цел алгоритъма изпълнява определен брой итерации. При всяка итерация се

изчислява логаритъм от вероятността зададеният обучаващ пример да задава текущото съвместно вероятностно разпределение. EM-алгоритъма се опитва да максимизира тази логаритмична вероятност. Итерациите спират, когато разликата между получените от две итерации логаритмични вероятности стане достатъчно малка (например от порядъка на  $10^{-4}$ ). Алгоритъмът не може да обучава условните разпределения за непрекъснати възли.

#### Адаптиране на базата на реализирани случаи

Адаптирането се извършва при структурно определена и начално обучена мрежа. То се извършва чрез последователно опресняване на вероятностните разпределения в БДМ на базата на наличния опит и чрез въвеждане и разпространяване на факти. Адаптиране на MAR.NET е необходимо да се извършва след всяко регистриране на инцидент или получаване на информация от проверка, изследване или наблюдение.

Опитът за даден дискретен възел се представя като множество от брой факти  $Alpha_0, \dots, Alpha_{n-1}$ , където  $n$  е броят конфигурации на родителите на възела.  $Alpha_i > 0$  за всяко  $i$ .  $Alpha_i$  означава броят пъти родителският възел да е бил в състояние на  $i^{ma}$  конфигурация. Броят е със значение на честота и е неотрицателно реално число.  $Alpha_i$  се съхранява в таблици на опита присвоени на възлите определени за обучение. Останалите възли се адаптират по правилата за разпространение на вероятности в БДМ.

#### Въвеждане на експертен опит

Понятието за експертен опит в бейзовите мрежи може да се представи като количествена памет, която може да базира както на количествена експертна оценка така и на станали случаи. Разпространяването на опита е свързано с процес на изчисляване на приорното условно разпределение за променливите в мрежата. Обработката на опита е свързана с процес на изчисляване на променените разпределения за параметрите, които определят условните разпределения на променливите в мрежата. Използвания за адаптиране на MAR.NET алгоритъм има възможност за управление на актуалността на усвоения опит посредством специални таблици редуциращи ефекта на миналото. С тяхна помощ експерта може да адаптира модела към промените в обекта и системата му за управление на безопасността. Това се постига чрез намаляване на опитния брой  $Alpha_i$  с редуциращ фактор  $Delta_i$ , който е неотрицателно реално число между 0 и 1 но обикновено близко до 1. Подробно математическо описание на алгоритъма е дадено в Spiegelhalter and Lauritzen, (1990).

#### Структурно обучение

Една изключително интересна възможност на БДМ е възможността да се извлече структура на мрежата от описани случаи. Данните се структурират както в таб. 7. Такива алгоритми за структурно обучение са описани в (Spirtes, C. Glymour and. Scheines (2000) и Pearl (2000)) и са известни като PC алгоритми. Те се базират на тестове за независимост на променливите в модела. Тестовата статистика, която е приблизително  $\chi^2$  разпределена допуска условна независимост. При големи стойности на

теста се отхвърля хипотезата за независимост. Нивото на значимост  $LC$  е стойност между 0 и 1 определя вероятността за отхвърляне на вярна хипотеза за независимост. При увеличаване на  $LC$  ще се получат повече условни вероятностни връзки между променливите в модела. Препоръчаната стойност на  $LC$  е 0.05.

При структурно обучение на MAR.NET с данни за 122 случая на регистрирани инциденти в р-к "Бабино" и  $LC=0.05$ , без въвеждане на друг опит са получени условни връзки между следните променливи: "Групи работници – професия" → "Час на възникване на инцидента", "Общ трудов стаж" → "Човешки фактор", "Степен на образование" → "Ден след последната почивка" → "Отклонение от нормалните действия". При увеличаване на  $LC$  на 0.1, се появява връзка между "Час на възникване на инцидента" → "Трудов стаж в предприятието". Тази възможност дава още една насока за евентуално преразглеждане на разбирането на експерта за механизма на проявление на рисковете в обекта и оттам на модела MAR.NET.

### Симулиране на случаи

Мощно средство за изследване на безопасността чрез модела MAR.NET е симулирането на случаи. Случаите се генерират във вида показан в таб. 7. Симулацията се базира на вариация на текущото приорно разпределение. Генерираните от симулацията случаи могат да се използват за обучение на мрежата с евентуални събития. Задаването на определен процент липсващи стойности в генерираните случаи имитира неопределеност в информацията. Симулирани случай е по-целесъобразно да се генерират от имитационни модели на работата на реални подсистеми от обекта. По този начин може да се тества устойчивостта на системата за безопасност като се наблюдава постериорното разпределение след адаптиране на модела със случаи, които до момента не са регистрирани в практиката но представляват ценна информация за управлението на безопасността.

Огромно предимство на модела е възможността да извлече от системата информация чрез разпространяване на вероятности срещу посоките на ребрата. Така може да се симулира увеличаване или намаляване на тежестта на инцидентите и да се проследи по мрежата кои конфигурации от състояния имат съществен дял в това. Така например ако мрежата е слабо обучена и вероятностните разпределения са близки до униформеното то икономическия ефект от изпълнението на програми за безопасност се оценява като по-лош от непредприемането на действия. С обучението на MAR.NET с данни само за 10 случая, картината започва да изглежда съвсем различно. На преден план излиза ефекта от приоритизирането предвидените програми.

### ЗАКЛЮЧЕНИЯ

Разработена е интелигентна система за поддържане на управлението по безопасност MAR.NET. Системата е настроена за спецификите на минния бранш но може да се адаптира и за други браншове. Началното обучение на

системата, се препоръчва да става отделно за възгледобива, отделно за рудодобива и отделно за нерудните полезни изкопаеми. Адаптирането от своя страна следва да се извършва за конкретните обекти върху обектно наследени копия на обучена система. Начинът на работата със MAR.NET е изключително интуитивен и ясен. Структурата на мрежата и свързаните с възлите ѝ състояния и вероятности са винаги пред погледа на специалиста. Няма скрити слоеве, които да затрудняват интерпретацията на извлечените от системата изводи. MAR.NET е в състояние да извлече заключения и при висока степен на неопределеност. Интуитивният начин на описание е комбиниран със строгия математически апарат на Бейзовата вероятност. Фактите и неопределеността разпространяват ефектите си в мрежата глобално. MAR.NET може да се обучава глобално за бранш и да се адаптира за спецификите на конкретен обект. Обект на адаптиране са както вероятностите на отделните състояния, така и структурата на мрежата. Това предоставя два пътя за отразяване на разбирането на експертите в модела. MAR.NET може да се обучава, както експертно, така и от файл с данни за регистрирани или симулирани случаи. Добре обучена мрежа на MAR.NET може да се използва за обучение на специалисти. Мрежите на MAR.NET адаптирани за различни обекти могат да комуникират по между си вкл. през Интернет, като образуват интелигентна мрежа от още по-високо ниво с разпределено изчисление и възможност за обмен на знания.

### ЛИТЕРАТУРА

- М. Михайлов, Т. Петров. Анализ на риска от злополуки в минните дружества за добив на въглища с MAR 2, МГУ, София 1997
- М. Михайлов, Причини за трудови злополуки, Изд. Къща "Св. Иван Рилски", София, 2001
- William W. Hines. Probability and Statistics in Engineering and Management Science, John Wiley & Sons, NY, Chichester, Brisbane, Toronto, 2000
- Е. С. Вентцель. Теория вероятностей. Москва, "Высшая школа", 2001
- Stephen J. Guastello. Chaos, Catastrophe and Human Affairs: Application of Nonlinear Dynamics to Work Organizations and Social Evolution, Lawrence Erlbaum Assoc. Publisher, September, 1995
- Isabelle Stengers, Ilya Prigogine. The End of Certainty: Time, Chaos and the New Laws of Nature, 240 p. Free Pr., August 1, 1997
- Т. Петров. Нелинейни изследвания на данни за индустриални инциденти. Международна конференция САИ, София 1999
- Р.И. Трухаев, И.С. Горшков. Факторный анализ в организационных системах. Москва "Радио и связь", 1985
- М. Михайлов и колектив. Отчетчет по дог. ФУТ-8:, София 2002
- R. G. Cowell and A. P. Dawid. Fast retraction of evidence in a probabilistic expert system. *Statistics and Computing*, 2:37-40, 1992],

- S. L. Lauritzen. The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis*, 19:191-201, 1995
- D. J. Spiegelhalter and S. L. Lauritzen. Sequential updating of conditional probabilities on directed graphical structures. *Networks*, 20(5):579-605, Special Issue on Influence Diagrams, Aug. 1990
- P. Spirtes, C. Glymour and R. Scheines. Causation, Prediction, and Search. *MIT Press*, Adaptive Computation and Machine Learning, second edition, 2000
- J. Pearl. Causality: Models, Reasoning, and Inference. *Cambridge University Press*, UK, 2000

Препоръчана за публикуване от катедра  
"Руднична вентилация и техническа безопасност", МТФ



# INTELLIGENT COMPUTER SYSTEM FOR CONTINUOUS RISK EVALUATION AND DECISION SUPPORT OF SAFETY MANAGEMENT IN MINING

Todor Petrov

University of Mining and Geology "St. Ivan Rilski"  
Sofia 1700, Bulgaria  
E-mail: [tpp@mgu.bg](mailto:tpp@mgu.bg)

## ABSTRACT

The structure and the theoretical basis of Bayesian network - Mine Accident Risk dot Net (MAR.NET) for decision support in mining safety are presented. The network is composed from 22 nodes described with specific states designed for description of mining risk factors. In the heart of the network is conditional probability distribution of the chance node "Type of the accident" containing 37800 state configurations. A new formula for evaluation of safety risk level is proposed. For the purpose of decision making the network is extended to influence diagram. The possibilities of learning and adoption of MAR.NET to specific object in uncertainty and some application of simulated safety cases are discussed.

## INTRODUCTION

The information support of the safety management required processing of a large amount of data both of quantity and quality type. It is well known from the practice that taking into account only quantifiers of safety risk like coefficients and indexes of frequency and weight are not sufficient for characterizing and control of safety level. The quality system for risk evaluation is needed. The new generic ISO 18000 series devoted to quality safety management is an eloquent fact about the importance of the problem.

Today the investigation and registering of an accident are documented in minimum 60 data fields of different formats. The scrutinize of safety risks in the large mining companies ordinary includes more than 3000 massive of data for description of 50 and more accidents per year. Every one accident is classified in 21 indicators (tab. 1) any of which described with 2 up to 26 characteristics (states) (Michaylov and Petrov (1997), Michaylov and collective (2002)).

The psychology and cognitive science are ascertain the fact that the human mind cannot effectively manipulate such kind of data structures and meet serious difficulties to make an inference when the possible decisions have more than 3 alternatives. Such informational overload of the consciousness leads to ignorance of information and heuristic deciding of variants. The risk of bad decisions increasing and the safety become pursuit rather than achieved purpose. The problem is of particular interest in time critical decision-making.

Another important feature of safety management is the probability and fuzzy uncertainty attending decision-making. In practice this mean that the expected effects of actions for risk reducing also will be connected with some likelihood. What is the degree of certainty that one of the expected effects is most likely is the correct question. Rendering the expert opinions is everyday activity in safety management. The successive implementing of acquired data and expert experience in decision-making is of vital importance.

A typical approach for safety risks investigation by separately studying of the impact of isolated factor groups inevitably relevant to loses of information about the mutuality in the safety system. In terms of information such a "disjoint" is irreversible process. In other words the model we use to support our decisions those not reflect the real world because the interdependences in the safety system are very diligence removed since on methodological level.

New synergetic approach should apply for the purposes of risk investigation and decision support for improvement of safety. A model representing the mutual influence of dangers, human and control over the safety is needed.

The structure of intelligent computer system MAR.NET (My Accident Risk dot Net) for information fusion of databases and expert opinion is presented. The system is designed for practical use from safety managers in mining companies.

## INDICATORS OF SAFETY RISKS IN MAR.NET

In the best practice exist pursuit of an integrated system of indicators characterizing safety risks. An integrated system of safety indicators for mining industry was developed in University of Mining and Geology "St. Ivan Rilsky" in the beginning of 90 and last updated in Michaylov and collective (2002) (tab. 1). Every indicator has a set of characteristics structured in hierarchical groups. The number of groups in stets are between 2 and 26. Because of lack of space the characteristic sets are not present in this paper.

The defined indicators in table 1 can be used for quality investigation of safety risks in any other industrial branch. The characteristic sets of some of the indicators will be different. But the principle of studying and data manipulation remains the same. This is great advantage for implementation in practice and for software development.

The shown indicator system can be used for risk investigation in all industries. The worldwide practice shows that successive computer systems using artificial intelligence methodologies are developed for a local domain irrespective of some universality. MAR.NET is developed for investigation of safety risks in mining industry. The sub domains in knowledge base for mining branches - coal mining, metal and nonmetal mining both for underground and open pit, required different sets of characteristics for some indicators.

Table 1. Safety Risk Indicators.

Name	Short label
Time of occurrence	01.Hour
Occupation groups	02.Occupation
Degree of education	03.Education
Length of service	04.Practice
Length of service in entertainment	05.Practice_Co

Length of service in profession	06.Practice_Pro
Day after last rest (weekend)	07.Day_after
Hours after start of job	08.Hour_after
Place of accident	09.Place
Kind of job during the incident	10.Job
Kind of incident leading the accident	11.Incident
Human factor in cause of accident	12.Human_Factor
Material factor in cause of accident	13.Material_Factor
The dangers of the environment	14.Environment
Deviation from ordinary actions and conditions	15.1.Deviation_A 15.2.Deviation_C
Severity of the accident	16.Severity
Harmed parts of body	17.Body
Kind of injury	18.Injury
Period of health restore	19.Recover_Period
Safety precautions (risk reducing measures)	20.Measure
Machines related with the accident	21.Machinery

Place of accident is a typical indicator for which the characteristic set needs to be overwritten for the different mining objects. On the other hand the learning of MAR.NET system will be much more adequate for specific branch and adoption – for specific objects. The convergence of the systems is the next step.

#### DRAWING OF INFERENCE FOR SAFETY RISKS LEVEL

Reporting the safety risk level is the important end result of risk evaluation. It is hard to define all the different aspects of risks in notion of one safety level. For the purpose the following definitions are accepted:

The drawing of inference for safety risk level is a process of statistical conclusion for synergetic influence of the risk factors on the accident severity in uncertainty. The risk factors under review are shown in table 1.

In the terms of safety a classical definition for risk is the production:

$$RISK = PROBABILITY \times CONSEQUENCES$$

In the description of safety level all risks must be take into account. After execution of safety programs the object of evaluation is the remaining (current) risk. -  $R_c$ . The level of safety can be useful quantificator for comparison of objects and branches with one value. But the dimension of risk is specified from the dimension of consequences. Usually the consequences are classified as economical and human and social. In capacity of quantitative link can be used the count of loosed working hours. It is not necessary to be human-hours. The indicator 19.Recover\_Period described the consequences of the accidents with 10 discrete intervals – “A. Up to 3 days”, “B. 4 to 17 days”, ... , “J. 6000 days (means irrecoverable accidents)”.

The following expression for calculating the safety risk level are proposed:

If we accept that the threshold for sensitivity of risk evaluation is in probability of  $1/10^8$  and 3 loosed working days as a consequence, than the minimal safety risk is evaluated on  $R_0 = 3_x 10^{-8}$ . In that case the level of safety risk  $L_s$  can be calculated as a function of current risk -  $R_c$  and the threshold risk -  $R_0$  by (1).

$$L_s = \log(R_c / R_0) \tag{1}$$

The  $L_s$  posses some properties which makes it useful for calculating of risk level. First, since the risks are always positive quantity and  $R_c \geq R_0 > 0$  the value of level will be always calculable and  $L_s > 0$ . Second, when the current risk aligns with the threshold  $R_c = R_0$ , the safety risks level  $L_s = 0$ . And last but not least the human perceptions are determined from logarithmic levels as stated of generalized psychophysical law of Veber-Fehner. Take into these considerations the 10<sup>th</sup> basis of logarithm is recommended.

Besides the one-value quantification of safety risk the management is needed of detailed quality investigation based on the available knowledge. The data collection for risks and safety level of a specific object is made by exception (registry of accidents, failures, protocols of inspections etc.). It leads inexorable to statistical evaluation of excerpted conclusion for the real state of the safety system, which obviously is richer of properties (Вентцель (2001)). Nonlinear dynamics of manifestation of the incidents with possibility the safety system to pass in chaotic regime (Guastello (1997); Stengers and Prigogine (1997); Petrov (1999)) puts the question outside of the application range of well-developed methodologies for reducing of the problem dimension like deterministic factor analysis and classical statistical averages.

(Трухаев, Горшков (1985)).

#### Bayesian approach for statistical inference

The frequency interpretation of probability is called objective ore classical point of view in statistic theory. In the statistical decision theory the Bayesian approach is used to draw of conclusions in uncertainty. The approach offers a different interpretation of probability called subjective point of view. The idea of conditional probability takes a main place in Bayesian approach. From the investigator is required to use subjective probability as a measure of belief for the state of observed object (Hines, (2000)). This is more intuitive perception of probability, which means rather than chance than frequency. The level of belief is specified with the probability distribution for a given unknown parameter. This procedure is completely different of any other statistical approaches, where the uncertain parameters are treated as unknown constants. The Bayes approach required from investigator to think for unknown parameters as random variables.

#### The low for complete probability and Bayes Theorem

The evaluation of impact of co incidents on the accidents is in the base of detailed risk investigation. The mathematical fundament of the MAR.NET model is based on the following major dependencies:

Let  $A_1, A_2, \dots$  is an enumerated collection of events sharing a space of realization –  $S$ . The events in collection are mutually

independent with union S. Let B is another event. If the probabilities  $P(A_i)$  and  $P(B | A_i)$  are known for  $i \in I$  (I is the set of indexes of events) than it can be shown that:

$$P(B) = P(B | A_1)P(A_1) + P(B | A_2)P(A_2) + \dots, \quad (2)$$

a result known as law for complete probability;

$$P(A_j | B) = \frac{P(B | A_j)P(A_j)}{P(B | A_1)P(A_1) + P(B | A_2)P(A_2) + \dots}, \quad (3)$$

a result known as Bayes Theorem and;

$$P(A_1, A_2, \dots, A_n) = P(A_1 | A_2, A_3, \dots, A_n)P(A_2 | A_3, A_4, \dots, A_n) \dots P(A_{n-1} | A_n)P(A_n), \quad (4)$$

a result known as chain rule, with significant importance in Bayesian networks.

### Bayesian belief network - BBN

The Bayesian network is presented with directed acyclic graph with the following elements: the chance nodes representing random variables and the edges – probability independencies between the variables. The nodes have conditional probability tables assigned to describe the independencies. The nodes without the parent have unconditional (marginal) distribution. One of most powerful feature of Bayesian network is the global treatment of local uncertainty. In other words – the changes in probability distribution in one chance node are propagated to all the nodes linked with edges following the directions in the network. The propagation of probability against the directions is possible due to Bayes Theorem. For statistical description of net and propagation of probability the results (2, 3 and 4) are used.

Unlike of classical statistical inferences (which work rather than with confidence intervals than statement of probability) is, that the Bayesian inference completely described the fact, that the expectation alone cannot predict the probability of unexpected events. Prior information for unexpected events is needed. The necessity of prior information is the key part of Bayesian inference. Of course, this requirement is a weakness. Not always is easy to obtain the prior information, except of experts. But in lack of data for safety in an entertainment (newly created or thinly proficient) the opportunity to use experience from similar objects and from objects with complicated behavior of safety system, even not large, the advantages of such technology will stay clear for a short time.

### Drawing of inference in Bayesian network

Drawing of inference or making conclusions in Bayesian networks means calculating of conditional probability for some variables to be given information from the others. It is easy when all indications are lead from predecessors (parents) to child variables (nodes) of interests (B). But when an indication is given from child to parent node, the network must draw a conclusion against the direction of edges. The Bayes Theorem (3) is used for such a back propagation of probability.

### Decision-making and Influence diagram

Computer models for decision support can be developed on the base of pure BBN, but the conceptions of utility functions and the decisions are not clearly formulated and fully cover. The extension of BBN with two types of nodes represents an influence diagram (fig. 1). The influence diagrams are used for evaluation of different variants of decisions by calculating of expected utility of launched actions.

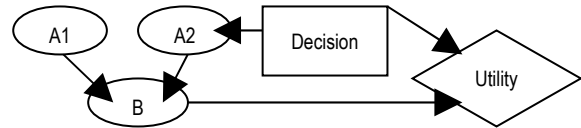


Figure 1. An example of Bayesian network extended to influence diagram with node "Decision" and node "Utility"

The decision nodes must be linked in chain in the logical consequence of dependency between variables in the model. The nodes from which the decision is depend, must be with known state before the decision is made.

### STRUCTURE OF MAR.NET

MAR.NET is designed as Bayesian belief network extended to influence diagram (fig. 2). The main purpose of MAR.NET is to support decisions for increasing of safety level in industrial objects, entertainment and whole branch. Currently the states of the nodes are adapted for the specific of coal mining. The nodes in the model correspond with the indicator variables described in tab.1. The states of indicators are labeled like the characteristics of the indicators. The conditional dependency between the variables can directly be read from the graph on fig. 2. The probability of different configurations of states in the net described the subjective point of view to happen an accident according to conditions given by states of parent nodes.

Marginal nodes in the root of the net are 14.Environment, 10.Job and 02.Occupation. The initial state of MAR.NET are uniform distribution of probability of the state configurations. The zero probability assigned to a state or configuration of state means striking off the possibility of occurrence of this state. The probability tables assigned to the marginal nodes are shown in table 2.

Social-human severity of accidents is evaluated of 16.Severity node. Economical risk of accident connected with 19.Recover\_Period; evaluated from utility node "Loses". The risk measured in missed working days is evaluated from chance node 19.Recover\_Period. All the 3 nodes are child of parent nodes and have conditional probability distributions.

Table 2. Probability table for node 10.Job

Material State	Probability
A. Transport and load	0.2
B. Ordinary exploitation	0.2
...	...
E. Other	0.2

The impact of accident is evaluated by nodes 01.However, 07.Day after and 08.However\_after (tab. 1) are object of additional study and are not connected in the MAR.NET at the present stage of development. The rest of the nodes have conditional probability distributions of their states (tab. 3).

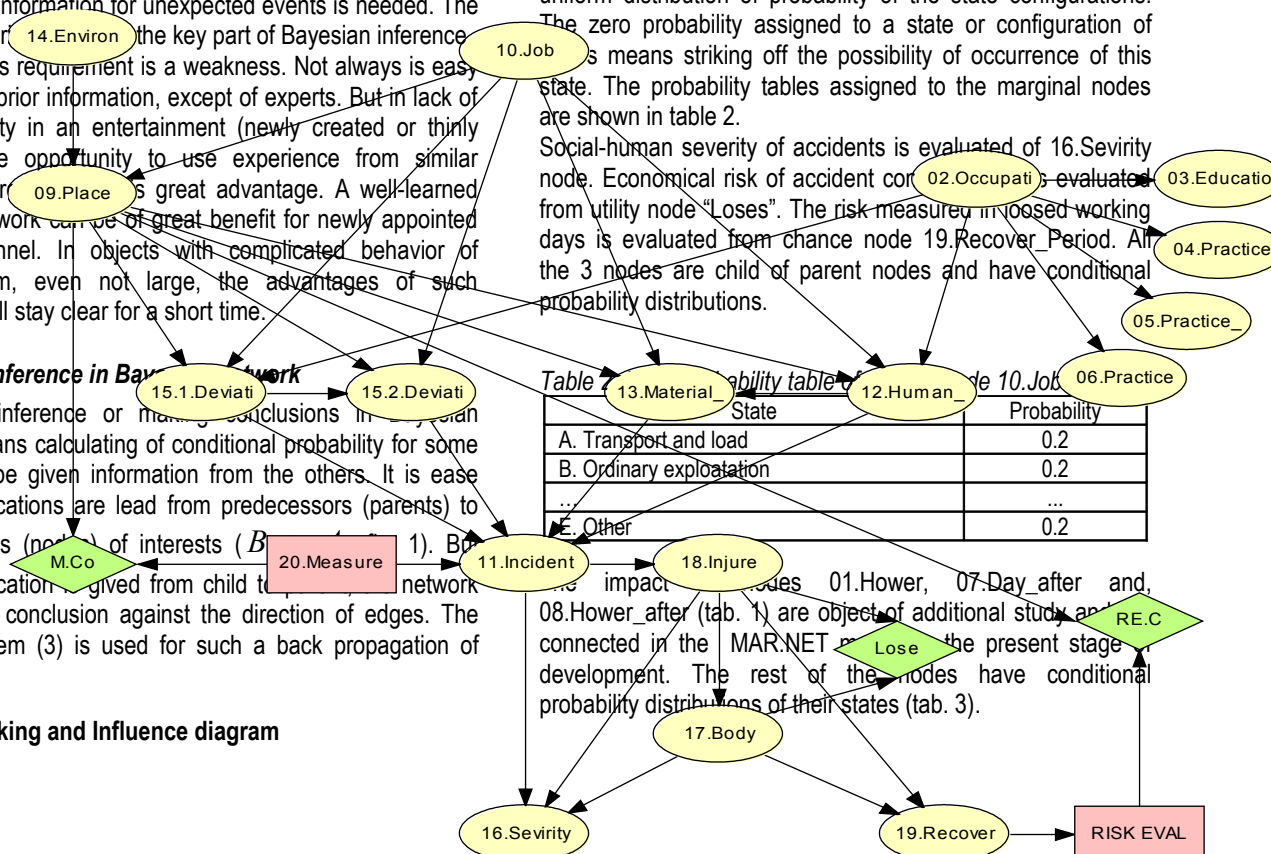


Figure 2. MAR.NET decision support system for improvement of safety level

Table 3. Initial conditional probability table  $P(17.Body|18.Injure)$

18.Injure \ 17.Body	A	B	...	Z
A. Head	0.25	0.25	0.25	0.25
B. Hands	0.25	0.25	0.25	0.25
C. Legs	0.25	0.25	0.25	0.25
D. Others	0.25	0.25	0.25	0.25
Total	1	1	1	1

Every state configuration (column) is an independent group with complete probability of 1.

In the heart of the MAR.NET is a chance node 11.Incident – “Kind of incident leading the accident”. In current realization the node is a child of five predecessors – four chance nodes and one decision node respectively 12,13, 15.1, 15.2 и 20 (see tab.1 for full names). The conditional probability table of node 11.Incident is a multidimensional massive given by probability distribution (5).

$$P(11.Incident|12.Human\_Factor,13.Material\_Factor, 15.1.Deviation\_Environment,15.2.Deviatin\_Action, (5) 20.Measure)$$

The massive dimension is  $18*14*15*5*2 = 37800$ . It is clear why for the experts is impossible to take into account all known configurations of conditional states.

The losses caused by the accident are described in utility node “Loses” (tab. 4)

The decision node 20.Measure – “Safety precautions (risk reducing measures)” is used for continuous evaluation of actions provided in safety programs.

Table 4. Utility table of node “Loses”

Body	A			B ...	
Injure	A	...	P	A	...
Loses	-90	...	-100000	-150	...

The effects of actions are propagated in MAR.NET through chance node 11.Incident. A simple question can be given by defining of two states of decision node 20.Measure: Action 0 and Safety Program. It means to do nothing or to execute a safety program. The cost of actions are specified in utility node “M.Cost” (tab. 5). The conditional independency of the nodes 09.Place and 20.Measure can be read of the graph in fig. 2.

Table 5. Utility table of the node “M.Cost”

Place	A		B		C
Measure	Action 0	Safety Program	Action 0	Safety Program	...
M.Cost	0	-5000	0	-10000	...

With the decision node “RISK EVALUATION” the expected utility of risk evaluation procedures are evaluated. For example if two state “Yes” and “No” are assigned as the decision the utility function will calculate expected utility of both actions. The pressure for starting of risk evaluation renders the increasing of loosed working days – node 19. The expenses related with the procedures of evaluation are given from utility function by node “RE.Cost” as conditional distribution determined of predecessor nodes “RISK EVALUATION” и “09.Place” (see fig. 2 and tab. 6).

The choice of alternative decision is make on principle of maximal expected utility. The global utility function  $U$  is a total of all expected local utility (5).

Table 6. Utility table of the node “RE.Cost”

Place	A		B		C
Risk Evaluation	YES	NO	YES	NO	...
RE.Cost	-1000	0	-2500	0	...

$$U = \sum_i u_i \tag{5}$$

$$u_i = \sum_j p_j u_j \tag{6}$$

where  $p_j$  is the conditional probability for occurrence of state configuration  $c_j$ , and  $u_j$  - is the value of utility related with this realization. For example the distribution of utility node Loses contains  $j = 104$  probability. But in calculation of  $U$  the local utility expected in nodes RE.Cost and M.Cost are taking into account, i.e.  $i = 3$ .

### INFERENCE OF SAFETY LEVEL WITH MAR.NET

If a new prior information is available the procedure for statistical inference must be started in order to update the probability distributions in the net. The propagation of known probability according to the independency between the variable (given by the edges) is executed. The posterior probability distribution for all the nodes is as a result. The expert can see the probability of any configuration of the node states and evaluated utility of alternative decision actions.

An example for node 10.Job - “ Kind of job during the incident” is given on figure 3.

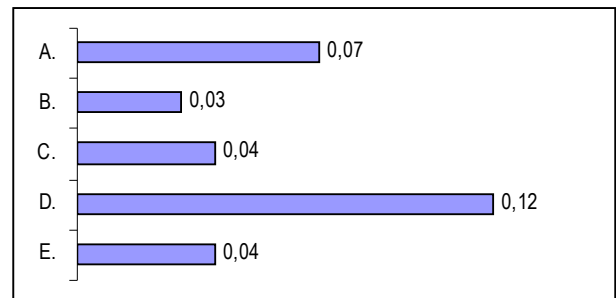


Figure 3. Posterior probability distribution of node 10.Job.

### LEARNINIG OF MAR.NET

The learning of the model can be done by 6 different ways. The first – imputing the known probability by hand for given configurations of states. According to dimensions of distributions this is a hard task. For the purpose of machine learning the special algorithms are developed. When the database for safety in the object is accessible, union of SQL queries can prepare the initialization of probability distributions. For the purpose of learning, the structure of MAR.NET is described in manner of (5) from the roots of network to the end child nodes. The SQL queries in the union follow described consequence. As a result the probability tables are formed.

### Learning from data cases

This type of learning is appropriate for initializing of probability distribution after structure defining of the net. A set of variables for which the prior information is available is specified. For any of the nodes related to the specified set the experience table is assigned. The experience tables count the number of realization of any specified configuration of states in the set. Learning set of variables (nodes) in MAR.NET envelope all indicators (tab. 7).

Table 7. Example of learning data cases for MAR.NET

N01	N02	N03	N04	N05	N06	...	N21
A	A	D	Q	A	B	...	N/A
C	I	D	D	N/A	C	...	C
...							

Where N01..N21 is the internal name of the nodes.

When there is no information about manifestation of some variable in the case the missing data is marked with "N/A" symbol. The codes A, B etc. corresponding with the labels of states of variables. The nodes represent the indicators (tab.1). The states are the characteristics of the indicators. The nodes and states of MAR.NET are compatible with databases of the software product for registering and analyzing of accidents Mine Accident Risk version 2002 (MAR). The product Mine Accident Risk are developed since 1995 in the department of "Mine Ventilation and Labour Safety" in MGU "St Ivan Rilsky". The learning method of MAR.NET is known as EM-algorithm commonly used in Bayesian network for graphical associated models with missing data (Cowell and Dawid, 1992; Lauritzen, 1995). The target of algorithm is enriching of conditional probability tables assigned to the nodes of network. For this purpose the algorithm performed a number of iteration. In any iteration logarithm of probability the given example to produce the current probability distribution is calculated. The EM-algorithm tries to maximize this log-probability. The iterations stop when the differences between log-probabilities obtained of two successive iterations became sufficiently small (for example of the order of  $10^{-4}$ ). The EM-algorithm cannot learn the conditional distributions for continuous nodes. In MAR.NET there are no continuous nodes.

### Learning adoption from data cases

Learning adoption of MAR.NET is necessary when a new accident is registered or new information from inspection, investigation or observation is available. The adoption of knowledge about the safety through consecutively updating of probability distributions in the net on the base of available experience is performed. The experience about a given discrete node is present as a set of counts for evidence  $Alpha_0, \dots, Alpha_{n-1}$ , where  $n$  is a count of configurations of the parent nodes.  $Alpha_i$  means the number of times a parent node to fail in  $i^{th}$  state configuration конфигурация. The count has a sense of frequency and is a nonnegative real number.  $Alpha_i$  is stored in experience table assigned of the nodes determined for learning adoption. The nodes for which there are no experiences are adopted according the rules of probability propagation in the net as discussed above.

### Entering expert opinions

The notion of experience in Bayesian networks can be introduced as a quantitative memory which can be based both on quantitative expert judgment and past cases. Dissemination

of experience refers to the process of computing prior conditional distributions for the variables in the network. Retrieval of experience refers to the process of computing updated distributions for the parameters that determine the conditional distributions for the variables in the network (Spiegelhalter and Lauritzen, (1990)).

The used in MAR.NET algorithm for entering the expert opinions allows control of the actuality of learned experience through special fading tables for reducing the impact of past. The fading factor  $Delta_i$  is used for reducing the experience count  $Alpha_i$ . The fading factor  $Delta_i$ , is a nonnegative real number between 0 and 1 but typically close to 1. The detailed description of the algorithm is given in Spiegelhalter and Lauritzen, (1990).

### Structure learning

A possibility to extract structure of the net from data cases is an exceptionally interesting feature of BBN. The data cases are structured in manner shown in tab. 7. The algorithms for structure learning of BBN are known as PC-algorithms (Spirtes, C. Glymour and Scheines (2000); Pearl (2000)). The independency tests for variables in the model is performed. The test statistic is approximately  $\chi^2$  distributed and allows conditional independency. The recommended value of level of confidence in which the zero hypothesis for independence is rejected is  $LC = 0.05$ .

Some interesting results were obtained in structure learning of MAR.NET with 122 data cases for registered accidents in coal mine of "Babino". The conditional dependency of following variables where accepted in  $LC = 0.05$ : Occupation  $\rightarrow$  Time of accident occurrence, Length of service  $\rightarrow$  Human factor, Education level  $\rightarrow$  Day after weekend  $\rightarrow$  Deviation from ordinary actions. In  $LC=0.1$  new dependence between Time of accident occurrence  $\rightarrow$  Length of service in entertainment is accepted".

The structure learning gives an alternative way the experts to reconsider his conceptions for safety in given object using artificial intelligence. When the understanding of safety risks manifestation is changed, the model of MAR.NET on structural level also can be changed.

### Simulation

Three approaches for obtaining simulated experience will be discussed. The first is by generating of simulated data cases and learning MAR.NET with EM-algorithm. The simulations are based on variations of the current prior distribution. The result of simulations must be in the format given in tab.7 in order to be useful from the learning algorithm. The simulation can be set to give a percent of missing data. The missing data imitate unknown probabilities for configuration of variable states and simulate uncertainty in the safety system. A more efficient approach is generating data cases by simulation models of real subsystems of the object. In both approaches the fixedness of safety system in case of occurrence of unregistered cases.

The third way of simulation is based on the powerful feature of Bayesian networks to derive conclusion against the direction of the edges. It can be simulate increasing of severity of accidents and after propagation of probability according Bayes Theorem to obtain posterior distribution of the predecessor nodes of interest.

## CONCLUSIONS

A system for decision support in mining safety MAR.NET is proposed. The system can be adopted for other branches saving the type nodes and the proposed structure. The states of the part the nodes can be different. It is recommended the learning of MAR.NET to realize on different copy of the system for open pit and underground coal mining and for metal and non-metal mining and quarries. Adoption of MAR.NET is adequate to realize for different objects on the learned instances of branch models.

There are not hidden layers in the MAR.NET. The structure is clear and ease to modify according to changes of expert opinions. The inference of safety level can be done in uncertainty, which is the usual case in safety management.

The well learned MAR.NET model can be used for education and training. The contemporary technologies allow the .NET models learned and adopted for different objects to communicate each other including via the Internet. Such a super-BBN in which the nodes are other BBNs can constitute intelligent net with distributed calculation and possibilities of knowledge exchange.

#### REFERENCES

- М. Михайлов, Т. Петров. Анализ на риска от злополуки в минните дружества за добив на въглища с MAR 2, МГУ, София 1997
- М. Михайлов, Причини за трудови злополуки, Изд. Къща "Св. Иван Рилски", София, 2001
- William W. Hines. Probability and Statistics in Engineering and Management Science, John Wiley & Sons, NY, Chichester, Brisbane, Toronto, 2000
- Е. С. Вентцель. Теория вероятностей. Москва, "Высшая школа", 2001
- Stephen J. Guastello. Chaos, Catastrophe and Human Affairs: Application of Nonlinear Dynamics to Work Organizations and Social Evolution, Lawrence Erlbaum Assoc. Publisher, September, 1995
- Isabelle Stengers, Ilya Prigogine. The End of Certainty: Time, Chaos and the New Laws of Nature, 240 p. Free Pr., August 1, 1997
- Т. Петров. Нелинейни изследвания на данни за индустриални инциденти. Международна конференция САИ, София 1999
- Р.И. Трухаев, И.С. Горшков. Факторный анализ в организационных системах. Москва "Радио и связь", 1985
- М. Михайлов и колектив. Отчетчет по дог. ФУТ-8:, София 2002
- R. G. Cowell and A. P. Dawid. Fast retraction of evidence in a probabilistic expert system. *Statistics and Computing*, 2:37-40, 1992],
- S. L. Lauritzen. The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis*, 19:191-201, 1995
- D. J. Spiegelhalter and S. L. Lauritzen. Sequential updating of conditional probabilities on directed graphical structures. *Networks*, 20(5):579-605, Special Issue on Influence Diagrams, Aug. 1990
- P. Spirtes, C. Glymour and R. Scheines. Causation, Prediction, and Search. *MIT Press*, Adaptive Computation and Machine Learning, second edition, 2000

*Recommended for publication by Department of*  
Mine ventilation and occupational safety, Faculty of Mining Technology