ВЛИЯНИЕТО НА АНРИ ПОАНКАРЕ ВЪРХУ ГЕОМЕТРИЯТА И ВРЪЗКИТЕ Й С ФИЗИЧНИЯ СВЯТ

СТАНЧО ГЕНЧЕВ ДИМИЕВ, РУМЯН ГЕОРГИЕВ ЛАЗОВ

IMPACT OF HENRY POINCARE ON THE GEOMETRY AN ITS RELATIONS WITH THE PHYSICAL WORLD

STANCHO GENCHEV DIMIEV, RUMYAN GUEORGIEV LAZOV

We examine parts of the Poincaré work related with geometry and its interconnections with the physical world. Comments on his philosophical views on the essence of geometry are given.

KEY WORDS: non-euclidean geometry, Lie groups, discrete groups, automorphic functions, conventionalism

Увод. Въпросът за произхода на геометричните истини излиза извън рамките на математиката. На философско ниво този въпрос се свързва с теорията на познанието на Емануил Кант, в която се твърди, че геометричните истини имат априорен характер, т. е. не зависят от опита, а са вкоренени в съзнанието на човека. По времето на Кант геометрията се е свеждала почти до тази, която днес се изучава в средното училище. По същество тази геометрия е създадена преди повече от две хилядолетия в антична Гърция, но знаменитите "Начала" на Евклид и до днес играят основна роля за педагогиката на математиката. Благодарение на многобройните коментатори на "Началата" една от аксиомите на Евклид добива завидна известност в научната общественост. Касае се за аксиомата за успоредните, или така наречения 5-ти постулат, чието съдържание се превръща в един основен проблем за математиката до 19 век включително.

Изглежда, че не философската позиция на Кант е засилила интереса към въпроса за валидността на 5-тия постулат. Сакери и Льожандр са хвърлили доста сили, търсейки доказателство на този постулат, независимо от мнението на споменатия философ.

В първата четвърт на 19 век отношението към въпросния постулат се изменя радикално. Нещо повече, трима математици от различни страни, независимо един от друг, стигат до убеждението, че е невъзможно той да бъде доказан и развиват геометрия с отрицанието му като аксиома. Това са Лобачевски, Боиай и Гаус, създателите на една нова геометрия, отлична от Евклидовата. Потенциално априоризмът на Кант, съгласно който Евклидовата геометрия е единствено достъпна за човешкия разум, е опроверган.

Приемането на различни геометрии изменя радикално разбирането ни за връзките на математиката с физичния свят. Задачата за характеризирането му чрез подходяща аксиоматична система става една от централните теми за математиците на 19 век. В това направление Анри Понкаре има съществени математически приноси, създавайки успоредно своя оригинална философия.

Квадратични геометрии. Посочената в увода проблема е засегната от Поанкаре в неговата публикация от 1887 г. под заглавие "Относно основните проблеми на геометрията". Ето как изглежда началото на увода към тази статия:

"В логиката е невъзможно да се извлече нещо от нищо; всяко доказателство предполага известни предпоставки. Ето защо, математическите науки се опират на известен брой положения, които не се доказват. Можем да ги наричаме аксиоми, хипотези или постулати, без да сме задължени да ги третираме като факти, извлечени от опита (експеримента), нито да схващаме като аналитични съждения, или, накрая, като априорни синтетични съждения, важното е, че те несъмнено съществуват".

Първият параграф на статията е озаглавен "Квадратични геометрии", в които се изброяват трите основни 2-мерни геометрии:

1^{o}. Евклидовата геометрия, в която сумата от ъглите на триъгълника е два прави ъгъла;

2^o. Римановата геометрия, в която тази сума е по-голяма от два прави ъгла;

3^o. Геометрията на Лобачевски, в която тази сума е по-малка от два прави ъгъла.

"Тези три геометрии се основават на едни и същи аксиоми с изключение на 5-тия постулат на Евклид, приет в първата и отхвърлен в останалите. Освен това принципът, съгласно който две точки напълно определят една права, прави изключение в геометрията на Риман и не допуска изключения в останалите две.

Ако се ограничим за две измерения, геометрията на Риман допуска много просто тълкуване: тя не се отличава, както е известно, от сферичната геометрия, стига за прави да приемем големите окръжности върху сферата."

Посоченото по-горе третиране на сферичната геометрия се имитира нататък така, че да стане възможно неговото разпространение и за геометрията на Лобачевски. За тази цел Поанкаре използва повърхнини от втора степен. Ако с **К** означим повърхнина от втора степен, уславяме се да наричаме *права* всяко равнинно диаметрално сечение на **К**, а всяко недиаметрално равнинно сечение - *окръжност*. Понятието дължина на отсечка се определя по естествен начин като дъга с определени краища лежаща върху права, така както я определихме по-горе, т.е. като част от равнинно диаметрално сечение. За да определи *дължина* на така взетата отсечка Поанкаре използва понятия от развитата до негово време проективна геометрия. Именно, двата края на отсечката и двете безкрайни точки на коничното сечение К определят своето двойно (анхармонично) отношение. Уславяме се под дължина на разглежданата отсечка да разбираме логаритъма от въпросното двойно отношение ако се касае за хипербола, и същия логаритъм умножен с комплексното число $-i$, ако имаме предвид елипса, в частност сфера. Понятието *ъгъл* между две пресичащи се прави се определя пак с помощта на двойното отношение, като за целта се използват двете допирателни и двете праволинейни образуващи в дадената пресечна точка, разбира се при подходящи условия за това дали се касае за хиперболоид или не.

Предвид на впечатляващата изразителност, нататък цитираме дословно текста:

"Между ъглите и дължините, както бяха определени, ще съществуват ред съотношения (релации), които съставят съвкупността на теоремите, аналогични на тези в равнинната геометрия (на Евклид). Тази съвкупност можем да наречем квадратична геометрия, защото като отправна точка използвахме повърхнини от втора степен".

Ако взетата основна повърхнина е елипсоид, получаваме геометрията на Риман, ако е двуполюсен хиперболоид – получената квадратична геометрия не се различва от тази на Лобачевски, ако е елиптичен параболоид – евклидовата геометрия. Разбира се, броят на квадратичните геометрии нараства ако разгледаме всевъзможните случаи на повърхнини от втора степен. Измежду тях главна роля играят тези повърхнини от втора степен, които имат център – посочените три геометрии.

Любопитно е, че Поанкаре поставя въпроса защо тези геометрии не са били забелязани толкова време – цели две хилядолетия! Отговорът се крие в предубежденията от различно естество, свързани с разбиранията за математика.

Ли-групи. Трябва да отбележим, че не по-малък интерес в излаганата по-горе статия на Поанкаре представлява доказателствената техника, която той използва. Това е теорията на така наречените днес групи на Ли, създадена от неговия съвременник норвежкият математик Софус Ли.

Поанкаре взема предвид, че и за трите основни геометрии групите на движенията (преместванията) им са транзитивни, а изотропните им групи са 1-мерни. Тук се налага да говорим на по-специалния съвременен език на Ли-гупите и Ли-алгебрите. Използвайки инфинитезимални методи на Ли, Поанкаре успява да класифицира 2-мерните хомогенни пространства на тримерни групи с точност до локална еквивалентност. Разполагайки с тази класификация той успява да различи няколкото възможни случая с помощта на някои геометрични свойства. В съвременен вид доказателствата в разглежданата статия се опират на формулата на Campbell за произведението на експоненциалите $\mathbf{e^X}$, $\mathbf{e^Y}$ на две "малки" векторни полета **X, Y,** т. е. $\mathbf{e^X.e^Y = e^Z}$, на теоремата на Поанкаре-Биркхоф-Вит, която осигурява единствеността на векторното поле $\mathbf{Z = \varphi(X,Y)}$, чиито формален ред е сходящ близо до $\mathbf{X = Y = 0}$, бидейки аналитична функция. Така се получава структура на Ли-алгебра. Посочената функция определя просто транзитивна група, което следва от свойствата на умножението в универсалната обвиваща алгебра. Нататък третата фундаментална теорема в теорията на Ли-групите дава възможност да се приложи смятането с резидууми, водещо до решаването на уравнение с частни производни.

Модели на неевклидовата геометрия, автоморфни функции. През втората половина на XIX век отношението на математиците към неевклидовите геометрии радикално се променя в сравнение с времената на Гаус, Лобачевски и Бойяй. Вероятно най-голяма заслуга за това имат Риман и Белтрами, който построява повърхнина в тримерното евклидово пространство, върху която се реализира (при подходящо определение на основните геометрични обекти) неевклидовата геометрия на Лобачевски-Бойяй. Въпреки силното впечатление от построението на Белтрами, този успех на неевклидовата геометрия оставя известна неудовлетвореност и чувство за половинчатост у математиците: върху псевдосферата на Белтрами неевклидовата геометрия се реализира само локално, т.е. построената геометрия върху повърхнината съответства на геометрията на част от неевклидовата равнина. Нещо повече, като цяло, топологично псевдосферата значително се отличава от неевклидовата равнина: там например съществуват затворени криви, които не могат с непрекъснато преобразуване да се свият в точка. Тези обстоятелства пораждат силен интерес към построението на Феликс Клайн, който реализира модел на *цялата неевклидова равнина* в единичния кръг на евклидовата равнина. Интересът към неевклидовите геометрии съпътства цялата дейност на Поанкаре и малко след Клайн той построява друг модел на неевклидовата равнина върху горната полуравнина на комплексната равнина. Понятията "точка", "права", "окръжност", "разстояние" от неевклидовата геометрия се дефинират с помощта на комплексната геометрия на равнината и по такъв начин се изгражда цялата геометрична картина на неевклидовия свят. Един от най-съдържателните и плодотворни аспекти на схващането на Поанкаре за геометрия въобще (и за конкретния модел в частност) е определението на понятието "движение". Именно това понятие, определено в термините на дробно-линейните трансформации на комплексната равнина, се поставя в основата на изучаването на конкретната геометрия. Особено съществен е фактът, че тези трансформации образуват група - това схващане е изцяло в духа на Ерлангенската програма на Клайн, която свързва изучаването на всяка геометрия със свойствата на определена група. Моделът на Поанкаре на равнинната неевклидова геометрия се оказва много плодотворен и подсказва обобщение в пространства с повече измерения - един аспект на геометричните изследвания започнат от Поанкаре и запазил важността и актуалността си до днес. Тези чисто геометрични изследвания биха поставили Поанкаре на почетно място сред геометрите на XIX век, но дълбочината на геометричните му идеи става ясна като вземем предвид връзките между неевклидовата геометрия и един дял от анализа. Става дума за задачата за интегриране на диференциални уравнения с алгебрични коефициенти. Изследванията на немския математик Л. Фукс в чисто аналитичен аспект възбуждат силен интерес и в Германия, и във Франция - много от активно работещите аналисти се включват в разработването на това ново поле, важно както в теоретичен, така и в приложен аспект. Първият фундаментален резултат на Поанкаре в тази област е: решенията на такива уравнения трябва да бъдат автоморфни функции, т.е. такива, които остават инвариантни при дробно-линейни трансформации на комплексната равнина. Появата на автоморфните функции прави предмета изключително актуален, тъй като аналогията с елиптичните функции - един от върховете на анализа на XIX век, е предизвикателство за аналистите. Подходът на Поанкаре обаче е радикално различен и изключително резултатен: той схваща елиптичните функции като инвариантни при определена група движения на евклидовата равнина. При този подход основна роля играе фундаментална област за дадена функция, която в елиптичния случай е успоредник - следователно изучаването на елиптичните функции естествено води до паркетирането на евклидовата равнина с успоредници. Следващата крачка е типична за методите на Поанкаре - той разглежда група движения на неевклидовата равнина, построява паркетиране с специфични фигури (които могат да се разглеждат като неевклидови аналози на успоредник) и оттук извежда основните свойства на автоморфните функции. Трябва да се отбележи, че в хода на тези изследвания се използва целия геометричен арсенал на неевклидовата геометрия, включително и многомерните неевклидови пространства, чиито групи от трансформации пряко се свързват с поведението на автоморфните функции. Може би тук трябва да отбележим, че връзките между автоморфните функции и многомерната неевклидова геометрия играят централна роля в изследванията по тази тематика и до днес.

Коментарии. Една опростена версия за философските разбирания на Кант за познанието на природата е следната: ние не познаваме и не можем да опознаем природата. В нея виждаме само това, което ни позволява разумът. Съгласието между Природата и математиката се създава от човешкия разум. Така физическото пространство е априорно определено, геометрията е евклидова, а

математиката има скромната цел да изучава пространството такова, каквото то е предопределено за нас.

Преди да изложим възгледите на Поанкаре за пространството нека напомним за две важни работи по геометрия предшестващи Поанкаре, а именно: статията на немския математик Бернхард Риман, ученик на Гаус, "Относно хипотезите, лежащи в основите на геометрията" (1866) и статията на немския физиолог, физик и математик Херман Хелмхолц "Относно фактите, лежащи в основите на геометрията" (1868).

Счита се, че тези две статии са основоположни за създаването на геометрията на пространството и даже на многомерната диференциална геометрия. Съгласно първата от тях, явяваща се обобщение на Гаусовата теория на повърхнините, за основа се взема понятието дължина на дъга ds^2, която локално за малки околности на фиксирана точка се изчислява съгласно Питагоровата теорема ($ds^2 = dx^2 + dy^2 + dz^2$). Съгласно втората – за основа се взема наличието на движения (премествания) в пространството, възможността за които се обуславя от съществуването на твърди тела в природата. Ясно е защо Риман говори за "хипотези", а Хелмхолц за "факти". Първият развива геометрични идеи, свързани с идеята за абстрактно пространство, а вторият – за идеи свързващи геометрията с физическото пространство. От изложеното по-горе личи, че възгледът на Поанкаре за пространството е различен. Както видяхме, Поанкаре счита, че не сме задължени да схващаме основните положения нито като хипотези, нито като факти.

В своята книга с популярна насоченост "Наука и хипотеза" той настоява, че въпросът валидна ли е евклидовата геометрия за физическото пространство е лишен от смисъл. При наличието на повече геометрии една от тях не може да бъде по-вярна от друга, би могла да бъде само по-подходяща. Понастоящем, евклидовата геометрия е, и ще остане, най-подходящата. Това е така защото:

1. Тя е най-простата, и то не само защото нашата ментална интуиция е пряко свързана с нея, а защото тя е проста сама по себе си.

2. Защото тя е достатъчно свързана с естествените твърди тела, които можем да сравняваме и да измерваме.

В това се изразява конвенционалността на геометрията. Това е конвенционалната философия на Поанкаре за пространството заключаваща, че геометрията на "нашия свят" е евклидова не защото е априорно мотивирана, а защото е най-удобна.

Трябва да отбележим, че в съвременната космология едромащабната пространствена структура се съгласува добре не с евклидовата геометрия, а с геометрията на Лобачевски. Съвременната обща теория на относителността разполага с достатъчно добри възможности за провеждане на измервания, потвърждаващи горното твърдение. Това се отнася за геометрията на това, което наричаме "пространство-време", което притежава физична интерпретация. Пространствено-времевата геометрия на специалната теория на относителността, развита от Поанкаре (1906) преди Айнщайн, също е неевклидова. Виждаме, че пространствено-времевата геометрия не се вписва удачно в конвенционалната философия за обичайното пространство. Метриката на Минковски със сигнатура (1,3) върху пространството-време налага съвсем различни свойства от тези в обичайната риманова геометрия с положително определена метрика.

ЛИТЕРАТУРА

1. Об основаниях геометрии, Сборник класических работ по геометрии Лобачевского и развитию ее идей, Москва, 1956
2. R. **Penrose**, *Physical space-time and nonrealizable CR-structures*. // Bulletin of the American Mathematical Society, vol. 8, No 3, May 1983
3. В. В. **Никулин**, И. Р. Шафаревич, *Геометрии и группы.* Москва, 1983:

ФУКСОВИ ГРУПИ И АВТОМОРФНИ ФУНКЦИИ У А. ПОАНКАРЕ

ПЕТЪРРУСЕВ

FUCHSIAN GROUPS AND AUTHOMORPHIC FUNCTIONS BY H. POINCARE

PETER RUSEV

A very short survey of the notions of a discontinuous group of analytical automorphisms and its fundamential region is given. A Fuchsian group is defined as a discontinuous group of linear transformations of the unit disk in the complex plane. The Poincare model of Lobachevsky's planimetry is described. The construction of an authomorphic function with respect to a Fuchsian group by means of Poincare's series is cleared.

KEY WORDS: fuchsian groups, authomorphic function

1. Прекъснати групи и автоморфни функции. За област $D \subset \mathbb{C}$ с $\mathbf{G}(D)$ означаваме групата на аналитичните автоморфизми на D. Ако G(D) е подгрупа на $\mathbf{G}(D)$ и z е точка от D, множеството $\mathcal{O}z := \{Tz : T \in G(D)\}$се нарича орбита на z относно G(D).

Групата $G \subset \mathbf{G}(D)$ се нарича прекъсната, ако $\mathcal{O}z$ няма точка на сгъстяване в D за всяко $z \in D$. $F \subset D$ се нарича фундаментална „област" на прекъснатата група $G \subset \mathbf{G}(D)$, ако сечението $F \cap \mathcal{O}z$ съдържа само една точка за всяко $z \in D$.

Мероморфна функция f, дефинирана в областта D, се нарича автоморфна относно прекъснатата група $G(D) \subset \mathbf{G}(D)$, ако за всяко $z \in D$, $f(\mathcal{O}z) = f(z)$, където $\mathcal{O}z$ е орбитата на точката z относно групата G(D).

Както е известно, $\mathbf{G}(\mathbb{C}) = \{z \to Tz = az + b,\ a \in \mathbb{C}^* = \mathbb{C} \setminus \{0\},\ b \in \mathbb{C}\}$. Ако $\mathbf{G}(\mathbb{C})$ е прекъсната подгрупа на $\mathbf{G}(\mathbb{C})$ и $T \in \mathbf{G}(\mathbb{C})$, то непременно $Tz = z + b$, $b \in \mathbb{C}$. Наистина, ако $a \neq 1$, съществува $\alpha \in \mathbb{C}$ такова, че $T\alpha = \alpha$ и, следователно, $Tz - \alpha = a(z - \alpha)$, откъдето получаваме че $T^n z - \alpha = a^n(z - \alpha)$, $n = 1,2,3,....$ Тогава, ако $|a| < 1$, от $T^n 0 = \alpha - a^n\alpha$ следва, че $\lim_{n\to\infty} T^n 0 = \alpha$. Ако $|a| = 1$, бихме получили че $|T^n 0| \leq 2|\alpha|$, $n = 1,2,3,....$ И в двата случая ще стигнем до извода, че редицата $\{T^n 0\}_{n=1}^{\infty}$ има поне една точка на сгъстяване, което не е възможно. Ако $|a| > 1$, до същия извод стигаме като имаме предвид, че T^{-1} принадлежи на групата $\mathbf{G}(\mathbb{C})$.

Добре известен факт е, че прекъснатите подгрупи на $\mathbf{G}(\mathbb{C})$ са следните:

$$\text{(I) } G(\mathbb{C}; \omega) = \{z \to z + m\omega,\ \omega \in \mathbb{C}^*,\ m \in Z\};$$

$$\text{(II) } \mathbf{G}(\mathbb{C}; \omega, \omega') = \{z \to z + m\omega + n\omega',\ \Im\left(\frac{\omega'}{\omega}\right) > 0,\ m, n \in Z\}.$$

В случая (I) фундаментална област е всяка ивица ортогонална на вектора 0ω и с ширина $|\omega|$, към която е причислена и една от ограничителните и прави линии. Автоморфна функция относно $G(\mathbb{C}; \omega)$ се нарича периодична и ω е нейния период. Ако $\omega = 2\pi$, примери за такива функции са $\sin z$ и $\cos z$. Ако $\omega = \pi$, такава е функцията $\tan z$.

В общия случай функция автоморфна относно $G(\mathbb{C}; \omega)$ има вида $f(z) =$

$$\sum_{k=-\infty}^{\infty} a_k \exp\left(\frac{2k\pi i}{\omega} z\right) \text{ или } h(z) = \frac{f(z)}{g(z)}, \text{ където } g(z) = \sum_{k=-\infty}^{\infty} b_k \exp\left(\frac{2k\pi i}{\omega} z\right)$$

Автоморфните функции относно групата G(|; ω, ω′) са тези мероморфни функции в |, за които ω и ω′ са периоди. Те носят названието елиптични функции. Пример за такава функция е Вайерщрасовата $\wp$-функция, която се дефинира чрез равенството

$$\wp(z) = \frac{1}{z^2} + \sum_{m,n \in Z}{}' \left(\frac{1}{(z - m\omega - n\omega')^2} - \frac{1}{(m\omega + n\omega')^2} \right)$$

Всяка елиптична функция с периоди ω и ω′ има вида $R_1(\wp(z)) + \wp'(z)R_2(\wp(z))$, където R_1 и R_2 са рационални функции.

2. Модел на Поанкаре на неевклидовата планиметрия на Лобачевски. За реализация на неевклидовата планиметрия на Лобачевски Поанкаре избира единичния кръг E в |, т.е. $E = \{z \in | : |z| < 1\}$. Групата Γ($E$) на движенията в този модел се формира от групата G(E) на аналитичните автоморфизми на E и изображението на E дефинирано чрез $z \mapsto \bar{z}$.

Известно е, че

$$G(E) = \{z \mapsto \lambda \frac{z - \alpha}{1 - \bar{\alpha} z}, |\lambda| = 1, |\alpha| < 1\},$$

т.е. групата Γ(E) се състои от дробно-линейните преобразувания на единичния кръг в себе си и от преобразуването $z \mapsto \bar{z}$. От геометричните свойства на тези преобразувания следва, че правите линии в модела на Поанкаре на неевклидовата планиметрия на Лобачевски са дъгите от окръжности, ортогонални на единичната окръжност в |, принадлежащи на E, както и диаметрите на E.

Разстояние между две различни точки z_1 и z_2 от "равнината" на Лобачевски в модела на Поанкаре, т.е. точки от E, се дефинира чрез равенството $d(z_1, z_2) = c\log(z_1, z_2, \alpha, \beta)$, $c > 0$, където α и β са пресечните точки на "носителя" на неевклидовата права линия, определена от точките z_1 и z_2, и единичната окръжност, а $(z_1, z_2, \alpha, \beta)$ е двойното отношение на комплексните числа z_1, z_2, α, β.

3. Фуксови групи. Тези групи са въведени от немския математик Лазар Фукс във връзка с негово изследване на функции дефинирани чрез "обръщане" на интеграли от решенията на линейни диференциални уравнения с рационални коефициенти.

По дефиниция фуксова група е прекъсната подгрупа на G(E). На Поанкаре дължим създаването на теорията на автоморфните функции относно фуксови групи. Поанкаре доказва, че всяка фундаментална област на фуксова група е многоъгълник в неговия модел на неевклидовата планиметрия на Лобачевски.

Както е известно, немалка част от математиците от 19-ти век са се отнасяли с "подозрение" към неевклидовата геометрия на Лобачевски дори и след фундаменталния труд на Бернхард Риман върху хипотезите, лежащи в основите на геометрията. Привличането на планиметрията на Лобачевски от Поанкаре в класическия комплексен анализ е изиграло може би решаваща роля за "легализирането" на идеите на Лобачевски и създадената от него неевклидова геометрия.

4. Автоморфни функции у Поанкаре. Поанкаре установява, че всяка фуксова група с

фундаментален многоъгълник с краен (четен) брой страни е изброима. Този факт е в основата на неговата теория и, по-конкретно, той създава възможността за дефиниране на съответните фуксови функции, т.е. на функции автоморфни относно фуксова група от посочения вид.

Ако G(E) е фуксова група, функцията F, мероморфна в E, се нарича автоморфна форма с размерност -k, k ∈ Z, ако каквато и да е линейната трансформация $z \mapsto Tz = \frac{az+b}{cz+d}$ от групата G(E), $F(Tz) = (cz + d)^k F(z)$ за всяко z ∈ E.

Нека R(z) е рационална функция, която няма полюси върху единичната окръжност и $\Gamma = \{T_n : z \mapsto (a_n z + b_n)(c_n z + d_n)\}_{n=1}^{\infty}$ е изброима фуксова група. Тогава, редът на Поанкаре $F(R; z) = \sum_{n=1}^{\infty} R(T_n z)(c_n z + d_n)^{-k}, k \geq 3,$ дефинира автоморфна форма относно групата Γ с размерност -k. Частното на две автоморфни форми относно една и съща изброима фуксова група и с една и съща размерност е автоморфна функция относно тази фуксова група. По този начин Поанкаре доказва съществуването на автоморфни функции относно изброима фуксова група.

5. Модулярни функции. Откритията на Поанкаре бързо намират отклик в математическите среди от края на 19-ти и началото на 20-ти век и най-вече в Германия. Обширни изследвания и резултати в областта на автоморфните функции дължим на Феликс Клайн, който наедно с това на Давид Хилберт е едно от последните големи имена от Математическия Институт на Университета в Гьотинген.

От края на 19-ти век датира и откриването на модулярните функции дължащо се на Рихард Дедекинд. Модулярна функция е функция автоморфна относно прекъсната подгрупа на групата G(H), където H е "горната" полуравнина, т.е. $H = \{z \in \mathbb{C} : \Im z > 0\}$.

За дефинирането на модулярна функция Дедекинд привлича групата $M(H) = \{\tau \mapsto \frac{a\tau + b}{c\tau + d}, a, b, c, d \in Z, ad - bc = 1\}$, която е наистина прекъсната подгрупа на G(H).

Ако ω и ω′ са такива, че $\Im\left(\frac{\omega}{\omega'}\right) > 0$, Вайершрас дефинира

$$g_2(\omega, \omega') = 60 + \sum_{m,n \in Z}{}' \frac{1}{(m\omega + n\omega')^4},$$

$$g_3(\omega, \omega') = 140 + \sum_{m,n \in Z}{}' \frac{1}{(m\omega + n\omega')^6}.$$

Тогава, ако $\Delta(\omega, \omega') = g_2^3(\omega, \omega') - 27 g_3^2(\omega, \omega')$, функцията

$$J(\tau) = \frac{\Delta(1, \tau)}{g_2^3(1, \tau)}, \Im\tau > 0,$$

както това установява Дедекинд, е автоморфна относно групата M(H).

ИЗСЛЕДВАНЕ ПОВЕДЕНИЕТО НА НЯКОИ КЛАСОВЕ ФУНКЦИИ С ПСЕВДОАСИМПТОТИ И АСИМПТОТИ С ПОМОЩТА НА СИСТЕМИ ЗА КОМПЮТЪРНА АЛГЕБРА

АННА ВЪЛКОВА ТОМОВА

REMARKS ON THE BEHAVIOURS OF SOME FUNCTIONS WITH PSEUDO ASYMPTOTES AND ASYMPTOTES

ANNA VALKOVA TOMOVA

We have defined the idea for pseudo asymptotes of differentiable functions. We proved one criterion for existence of pseudo asymptotes and asymptotes for differentiable functions. In this paper we restrict the attention over the behaviors of some functions with pseudo asymptotes and asymptotes. Using the system for computer algebra MATHEMATICA 4.0 we draw the graphics of some functions with pseudo asymptotes and asymptotes.

KEY WORDS: Definitions for pseudo asymptotes of differentiable functions, one criterion for existence of pseudo asymptotes and asymptotes for differentiable functions, formulae and graphics of some functions with pseudo asymptotes and asymptotes.

1. Въведение

Нека функцията $f(x)$,x $> x_0$ е дефинирана при x > x_0. Както е добре известно [1], необходимото и достатъчно условие да съществува асимптота за кривата $(c): y = f(x)$ при $x \to \infty$ е да съществува границата:

$$n = \lim_{x\to\infty}[f(x) - kx] \qquad (1)$$

при фиксирано к. Известно е също така, че при доказване на необходимостта на условието (1) к се намира по единствен начин:

$$k = \lim_{x\to\infty}\frac{f(x)}{x} \qquad (2)$$

В тази работа ще спрем вниманието си върху съществуването на функции с т.нар. "псевдоасимптоти", като най-напред дадем дефиниция на това понятие и след това ще изследваме аналитично и графично някои конкретни примери.

Понятието "псевдоасимптоти" за функция на една променлива е дефинирано в [2] и [3]. В [3] е разрешен пълно въпросът за съществуване на функции с псевдоасимптоти. Доказан е един критерий за съществуване на псевдоасимптоти и асимптоти за диференцируеми функции, който разделя два пъти непрекъснато диференцируемите функции на три групи: функции без асимптоти, функции с псевдоасимптоти и функции с асимптоти. В тази работа под формата на план и съдържание на практическо занятие посочваме конкретни примери за такива функции. Изследваме аналитично и илюстрираме графично поведението им спрямо техните псевдоасимптоти и асимптоти с помощта на системата за компютърна алгебра MATHEMATICA 4.0. Чрез тези примери се стремим да засилим вниманието на преподавателите по математика към използуване на системите за компютърна алгебра за получаване и илюстриране на сериозни математически резултати.

В [2] и [3] сме въвели следната дефиниция:

Дефиниция 1. *Нека функцията $f(x)$ е дефинирана при x > x_0.. Предполагаме, че границата (1) съществува, а границата (2) не съществува (в смисъл на крайна стойност). Тогава правата а с уравнение y(x) = kx ще наречем псевдоасимптота за функцията f(x) при $x \to \infty$.*

Забележка 1. Аналогично може да се дефинира псевдоасимптота при $x \to -\infty$.

Забележка 2. Лесно се доказва фактът, че 2 функции: $f_1(x)$ и $f_2(x) = f_1(x) + ax + b$, които се различават само в линейните си части, имат един и същ характер спрямо своите наклонени асимптоти, т.е. или

имат такива, или нямат. Същото се отнася и по отношение на по-горе дефинираните псевдоасимптоти. Не се изключва случаят, когато едната функция има хоризонтална асимптота или псевдоасимптота.

Забележка 3. Примери за функции с псевдоасимптоти са функциите lnx, ln(1+x) (псевдоасимптотите им са абсцисната ос) и др.

В [3] сме доказали следния

1.1. ***Критерий (достатъчно условие) за съществуване на псевдоасимптоти и асимптоти на диференцируеми функции.***

Теорема. *Нека* $f : R \to R$ *е два пъти непрекъснато диференцируема функция.Тогава:*

1)Ако $f''(x) \geq \frac{1}{1+x}, x \geq 0$, *функцията f(x) няма асимптоти при при* $x \to \infty$;

2)Ако $\frac{1}{1+x^{1+l}} \geq f''(x) \geq \frac{1}{1+x^2}, x \geq 1, 0<l<1$ *f(x) има псевдоасимптота,но няма асимптота;*

3) Ако $0 < f''(x) \leq \frac{1}{1+x^{2+l}}, l > 0, x \geq 0$ *f(x) има асимптота при* $x \to \infty$.

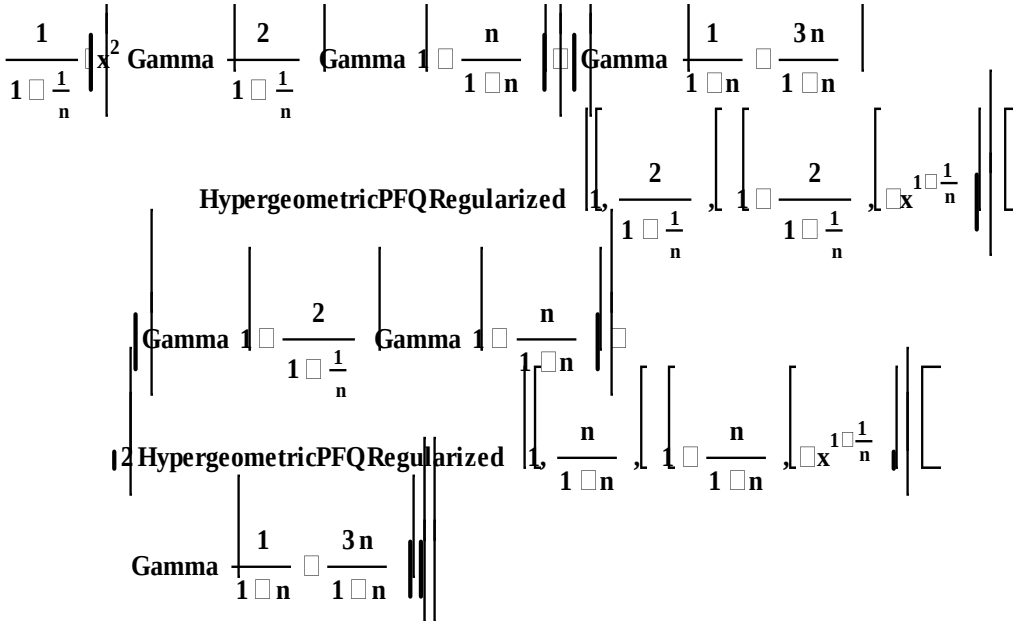
Забележка. Аналогично можем да формулираме и да докажем теорема за критерий (достатъчно условие) за съществуване на псевдоасимптоти и асимптоти на диференцируеми функции при $x \to -\infty$;

2. Класове функции с псевдоасимптоти и асимптоти.

Последователно интегрираме 2 пъти с помощта на системата за компютърна алгебра MATEMATICA 4.0:

```
Integrate[1/(1+x^(1+1/n)),x]
x/(-1-n)+x/(1+n)+x Hypergeometric2F1[n/(1+n),1,1+n/(1+n),-x^(1+1/n)]
Integrate[x/(-1-n)+x/(1+n)+x Hypergeometric2F1[n/(1+n),1,1+n/(1+n),-x^(1+1/n)],x]
```

В тази формула `Hypergeometric2F1[a, b, c, z]` е т.нар. хипергеометрична функция; тя се представя чрез следния ред: ${}_2F_1(a, b; c; \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!})$. Тази функция е решение на т.нар. хиперболично диференциално уравнение: $z(1-z)y'' + (c-(a+b+1)z)y' - aby = 0$. След второто интегриране получаваме следния клас функции с псевдоасимптоти:



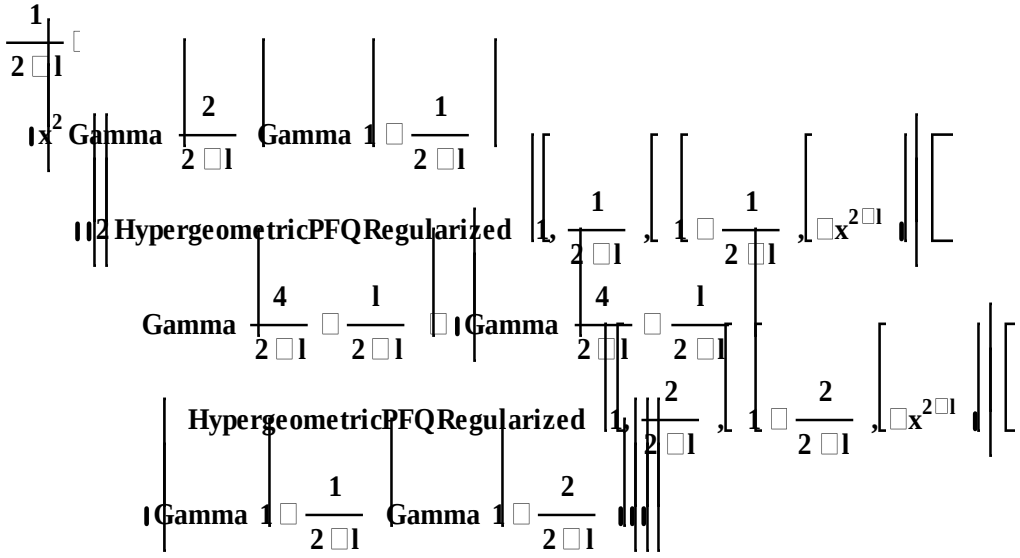
където: HypergeometricPFQRegularized[$a_1, \ldots, a_p, b_1, \ldots, b_q, z$] е т.нар. регуляризирана обобщена хипергеометрична функция ${}_pF_q(\mathbf{a}, \mathbf{b}; z)$. Функцията ${}_pF_q(\mathbf{a}, \mathbf{b}; z)$ се представя с развитието си в ред по следния начин: $\sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} \frac{z^k}{k!}$. С Gamma е означена известната гама – функция Γ(x).

2. 2. Функции с асимптоти

Последователно интегрираме 2 пъти с помощта на системата за компютърна алгебра MATEMATICA 4.0:

Интегрираме втори път :

Получаваме следния клас функции с асимптоти:



3. План за провеждане на практическо занятие по математически анализ на тема "Изследване поведението и построяване на графики на функции с пседоасимптоти и асимптоти с помощта на система за компютърна алгебра"

3.1. Теоретична част: обучаемите предварително са присъствували на лекция и са се запознали със съдържанието на темата и доказателството на критерия 1.1. Преподавателят обръща внимание на основните идеи и посочва съществуването на твърде широк клас елементарни и специални функции с псевдоасимптоти. Подчертава се сложността на пресмятанията и практическата полза за намирането на формулите с помощта на системи за компютърна алгебра.

3.2. Предлага се на обучаемите сами да намерят формули за функции с псевдоасимптоти и асимптоти.

3.3. Обучаемите построяват графики за получените функции в т. 3.2.

3.4. Правят се сравнения между графиките на функциите и се обясняват разликите в поведението им спрямо техните псевдоасимптоти или асимптоти.

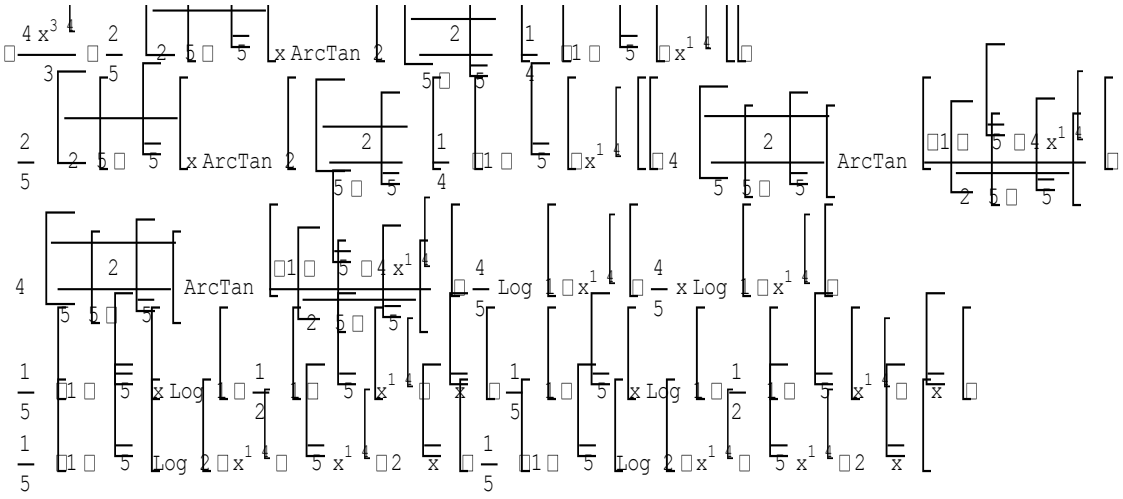
3.5. Ако се разполага с няколко системи за компютърна алгебра, правят се сравнения между резултатите, получени с тях при едни и същи задания.

3.6. Изводи: разсъждава се върху възможността за допущане на грешки и неточности при работа със системи за компютърна алгебра, констатирането и отстраняването им. Изтъкват се предимствата на работа със системи за компютърна алгебра за получаване и илюстриране на сериозни математически резултати.

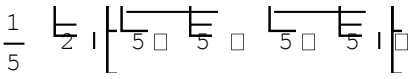
4. Графики на диференцируеми функции с псевдоасимптоти и асимптоти

4. 1. Графики на диференцируеми функции с псевдоасимптоти. Вариант: елементарни функции.

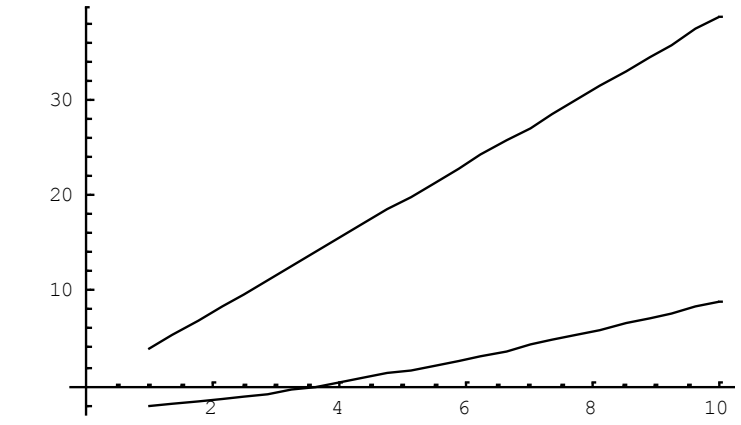
За частния случай: $f''(x) = \dfrac{1}{1+x^{1,25}}$, т. е. n = 4, след две интегрирания с помощта на системата за компютърна алгебра MATHEMATICA 4.0 получаваме формула за следната функция с псевдоасимптота:

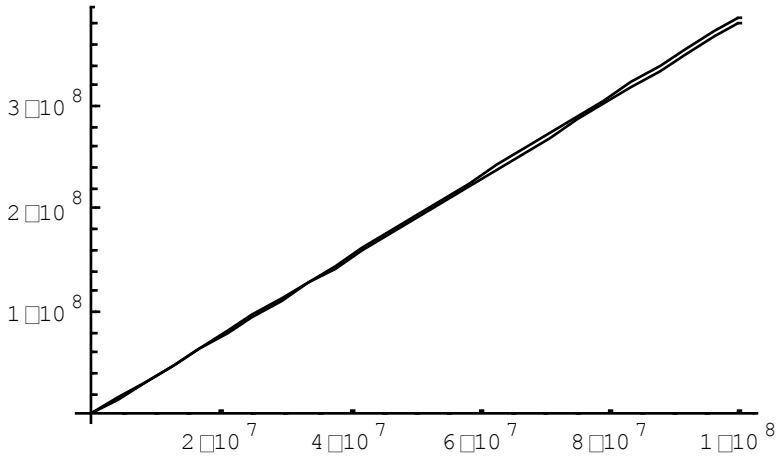


Намираме стойността на интересуващата ни граница (1): $k = \lim_{x\to\infty} \frac{f(x)}{x}$



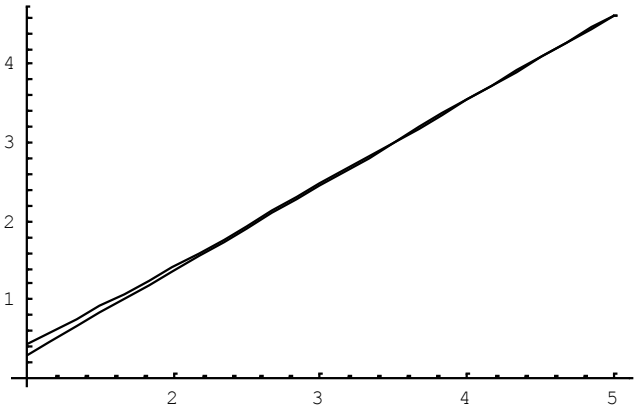
Чертаем графиката на функцията и на нейната псевдоасимптота за стойности на x в два интервала: [1,10], [1,100000000]. В последния случай графиката на функцията и графиката на нейната псевдоасимптотата са почти неразличими.





4.2. Графики за диференцируеми функции с асимптоти. Вариант: специални функции.

Разглеждаме случая: $f''(x) = \frac{1}{1+x^4}$,т.е. при l = 2. След две интегрирания с помощта на системата за компютърна алгебра MATHEMATICA 4.0 получаваме формулата за следната специална функция с асимптота и намираме двете граници (1) : $\frac{1}{2}\ p\Gamma\left[\frac{3}{4}\right]$ и (2): $-\frac{1}{2}\ p\Gamma\left[\frac{5}{4}\right]$. Построяваме графиките на функцията и нейната асимптота в интервала [1,5]:



5. Изводи

Сравняваме поведението на двете функции, разгледани в **4. 1** и **4. 2.** Както и очаквахме, едва при значителни стойности на аргумента графиката на функцията с псевдоасимтота и тази на нейната псевдоасимптотата са неразличими, докато при функцията с асимптота това се наблюдава при неколкократно по-малки стойности на x.

ЛИТЕРАТУРА

1. **Фихтенгольц** Г. М., Курс дифференциального и интегрального исчисления, Москва, 1975.
2 .**Томова** А., Бележки върху поведението на някои функции с псевдоасимптоти, Математически форум, ISSN 1311-297, Том 5, брой 2, март-април, 2003 г., стр. 55- 57.
3. **Томова** А., Един критерий за съществуване на псевдоасимптоти и асимптоти на диференцируеми функции. Бележки върху поведението на някои класове функции с псевдоасимптоти и асимптоти.

ВЪРХУ ПРИЕМНИТЕ ИЗПИТИ ПО МАТЕМАТИКА В НАЦИОНАЛНИЯ ПЕДАГОГИЧЕСКИ УНИВЕРСИТЕТ “М. П. ДРАГОМАНОВ”

Василий А. Швец

ON THE ENTERING EXAMS IN MATHEMATICS FOR THE NATIONAL PEDAGOGICAL UNIVERSITY “M. P. DRAGOMANOV”

Vasilii A. Shvec

The article says about measurements with the help of test of results of studying in mathematics with the pupils of secondary school in 1994 and 2005 years; the table of obtained data is given; the suggestions on improving of mathematical preparation of pupils are stated.

KEY WORDS: studying in mathematics, test results

Приемните изпити във висше учебно заведение са вълнуващо и значимо събитие в живота на всеки млад човек. През 2005 година приемните изпити по математика в НПУ “М. П. Драгоманов” бяха по-особени. Те се проведоха във формата на тест. Изпълниха го абитуриентите, които кандидатстваха за специалностите “математика и физика”, “математика и основи на икономиката” и “икономическа теория и информатика”. Тестирани бяха 228 кандидат-студенти. Това бяха предимно завършили средното училище с отлични оценки в дипломата по алгебра и основи на анализа и по геометрия. Повече от 60% от тях са носители на медали за отлично завършване на средното образование.

Тестът съдържаше 10 задачи от първо (средно), 10 задачи от второ (достатъчно) и 10 задачи от трето (високо) ниво на сложност. Правилно решена задача от първо ниво се оценяваше с 1 точка, от второ ниво – с 2 точки и от трето ниво – с 3 точки.

Задача от първо ниво се считаше за решена , ако беше подчертан като верен един от предложените отговори (А, Б, В, Г). В някои случаи освен избрания отговор абитуриентът трябваше да приведе и самото решение, на оставеното за тази цел място в теста.

За решаване на задачите от теста се даваха 180 минути. Лесно е да се пресметне, че за правилно решени всички задачи абитуриентът може да получи 60 точки.

Ето задачите от теста:

Задачи от първо ниво

1. Изчислете: 1,25+(3.7,9+0,75)

(А) 25**; (Б)** 25,7**; (В)** 26,7**; (Г)** *друг отговор***.**

2. Извършете действията $\frac{1}{a+1}+\frac{1}{a^2-1}$.

(А) $\frac{1}{a^2-1}$**; (Б)** $\frac{1}{a+1}$**; (В)** $\frac{a}{a^2-1}$**; (Г)** $\frac{2}{1+a}$

3. Намерете корените на уравнението: $|3x+0,6|=0,3$.

(А) $\{-0,1;0,3\}$**; (Б)** $\{-0,1;-0,3\}$**; (В)** $\{0,1;-0,3\}$**; (Г)** *друг отговор***.**

4. Опростете израза: tg $(5.\frac{\pi}{2} - \alpha)$.

(А) *tg* α**; (Б)** *-tg* α*;* **(В)** *ctg* α**; (Г)** *-ctg* α*.*

5. Намерете най-малкото цяло решение на неравенството: $|6 - x| < 5$.

(А) 0**; (Б)** 1**; (В)** 2**; (Г)** 3**.**

6. Функция е зададена с формулата $y = x \cos x$. Намерете стойността на функцията, която съответства на стойност на аргумента, който е равен на -π.

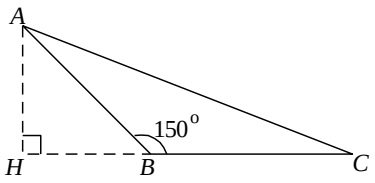
(А) 0**; (Б)** -π**; (В)** π**; (Г)** 1**.**

7. Намерете производната на функцията $y = 5e^x - x^2$.

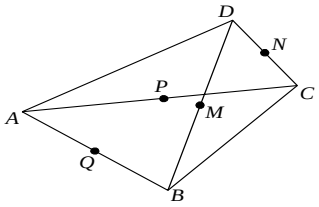
(А) $5e^x$**; (Б)** $-2x$**; (В)** $5-2x$**; (Г)** $5e^x - 2x$.

8. В равнобедрения триъгълник ABC (черт.1) $\angle ABC = 150^\circ$, $BC = 10$. Коя е дължината на височината AH на триъгълника?

(А) 5**; (Б)** 6**; (В)** 7**; (Г)** 8**.**



Черт.1



Черт. 2

9. Дадени са две кръстосани прави a и b. Колко равнини съществуват, които минават по правата a и са успоредни на правата b?

(А) 0**; (Б)** 1**; (В)** ∞**; (Г)** 1 или ∞**.**

10. Известно е, че равнината α е успоредна на правата b, а правата b е перпендикулярна на равнината φ. Кое е взаимното положение на равнините α и φ?

(А) *успоредни***; (Б)** *перпендикулярни***; (В)** *пресичат се под произволен ъгъл;* **(Г)** *отговор различен от* **(А); (Б); (В).**

Задачи от второ ниво

11. Теглото на брашното се отнася към теглото на зърното, от което е смляно, както 9:10. От площ 320 хектара е събрано по 45 центнера от хектар. Колко брашно ще се получи от това зърно?

(А) 12940 *ц* **;(Б)** 12950 *ц***; (В)** 12960 *ц***; (Г)** 12970 *ц.*

12. Извършете действията: $\left(\frac{x+y}{\sqrt{xy}}-2\right)\frac{\sqrt{x}}{\sqrt{x}-\sqrt{y}}$.

(А) $\frac{\sqrt{x}+\sqrt{y}}{\sqrt{xy}}$**; (Б)** $\frac{\sqrt{y}}{\sqrt{x}-\sqrt{y}}$**; (В)** $\frac{\sqrt{x}-\sqrt{y}}{\sqrt{y}}$**; (Г)** *друг отговор.*

13. Решете уравнението: $\lg\left(x^2+2x-7\right)-\lg\left(x-1\right)=0$

(А) -3**; (Б)** 2**; (В)** 1**; (Г)** 3**.**

14. Изчислете *sin* 75° . *cos* 15° без да използвате калкулатор или таблици.

(А) $\frac{1}{4}\left(2-\sqrt{3}\right)$**; (Б)** $\frac{1}{4}\left(2+\sqrt{3}\right)$**; (В)** $\frac{1}{2}\left(1+\sqrt{3}\right)$**; (Г)** $\frac{1}{2}\left(1-\sqrt{3}\right)$**.**

15. Решете неравенството : $\frac{2x-1}{x}\angle-\frac{1}{2}$.

(А) (-∞; 0)**; (Б)** (0; 0,4)**; (В)** (0,4**;** +∞)**; (Г)** *друг отговор.*

16. Намерете дефиниционното множество на функцията: $y=\frac{\lg\left(16-x^2\right)}{x+2}$

(А) (-4; 4)**; (Б)**(-4; -2)∪(-2**;** 3)**;(В)** (4**;** +∞)**; (Г)** $x\neq$-2.

17. Напишете уравнението на допирателната към графиката на функцията: $y=2\cos x$ в точката с абсциса $x_0=\frac{2\pi}{3}$.

(А) y=-$\sqrt{3}\,x$+1*;* **(Б)** y=-$\sqrt{3}\,x$+$\frac{2\pi\sqrt{3}}{3}$+1**;(В)** y=-$\sqrt{3}\,x$-$\frac{2\pi\sqrt{3}}{3}$*;* **(Г)** y= $\sqrt{3}\,x$-$\frac{2\pi\sqrt{3}}{3}$+1.

18. В кръг с лице 36π *см*^{2} е вписан правилен шестоъгълник. Изчислете лицето на този шестоъгълник.

(А) 56 *см*^{2}**; (Б)** 54$\sqrt{3}$ *см*^{2}**;(В)** 50$\sqrt{3}$ *см*^{2} **;(Г)** 80 *см*^{2}.

19. Върховете на триъгълника *ABC* имат координати *A*(2; 1; 3); *B*(2; 1; 5); *C*(0; 1; 1). Изчислете дължината на медианата *CM*.

(А) $\sqrt{12}$ **; (Б)** *2*$\sqrt{2}$ **;(В)** $\sqrt{13}$ **;(Г)** $\sqrt{10}$.

20. Точките *A*, *B*, *C* и *D* не лежат в една равнина (черт.2). Точките *M*, *N*, *P* и *Q* са среди съответно на отсечките *BD*, *CD*, *AC* и *AD*. Дължините на отсечките *AD* и *BC* са: *AD*=12, *BC*=14. Намерете периметъра на четириъгълника *MNPQ*.

(А) 24**; (Б)** 26**; (В)** 22 **; (Г)** 28.

Задачи от трето ниво

21. Докажете, че стойността на израза $\sqrt{3+2\sqrt{2}}-\sqrt{3-2\sqrt{2}}$ е естествено число.

22. Докажете тъждеството : $\sqrt{0,5+0,5\sqrt{0,5+0,5\cos 2\alpha}}=-\cos\frac{\alpha}{2}$, ако $\frac{3\pi}{2}<\alpha<2\pi$.

23. Решете системата: $\begin{cases}\sqrt{x}+\sqrt{y}=4\\ x+y=28\end{cases}$

24. Решете уравнението : $\cos 2x-\sin^2 x=0$.

25. Решете неравенството : $\sqrt{-2-x+x^2}\le x-1$.

26. Постройте графиката на функцията $y=\log_2(x-2)$.

27. Намерете интервалите на намаляване на функцията $y=4+3x^2-x^3$.

28. Докажете, че ъгълът, вписан в окръжност, е равен на половината от съответния му централен ъгъл.

29. Хорда от основата на конус има дължина a и съответна дъга α . Отсечката с краища върха на конуса и средата на хордата, е наклонена към равнината на основата под ъгъл β. Намерете обема на конуса.

30. В основата на пирамида лежи равнобедрен триъгълник с бедро b и ъгъл α при върха. Всички двустенни ъгли при ръбовете на основата на пирамидата са равни на φ. Намерете лицето на околната повърхнина на пирамидата.

Изложеното съдържание показва, че задачите:

а) № 1, № 11, № 21 представят съдържателна линия на училищния курс по математика – "Числа и действия с тях";

б) № 2, № 4, № 22, №24 – "Изрази и преобразования с тях";

в) № 3, № 5, № 13, № 14, № 15, № 23, № 25 – "Уравнения, неравенства и системи от тях";

г) № 6, № 16, № 26 – "Функции и техните графики"

д) № 7, № 17, № 27 – "Увод в диференциалното и интегралното смятане";

е) № 8, № 18, № 29, № 30 – "Геометрични величини – измерване и изчисление";

ж) № 9, № 19, № 28 – "Геометрични фигури и техните свойства";

з) № 10, № 20 – "Координати и вектори".

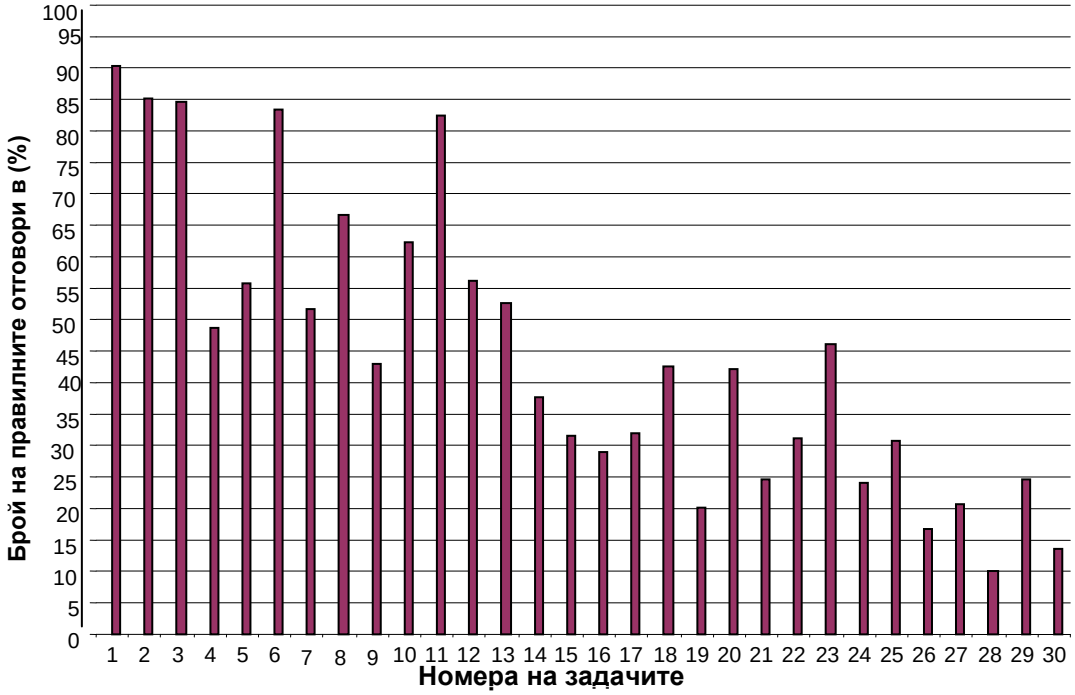
В текста не са включени задачи от съдържателните линии "Елементи на теорията на множествата, "Комбинаторика" и "Увод в теорията на вероятностите и елементи на статистиката", понеже тези съдържателни линии не са включени в програмите по математика за постъпване във висшите учебни заведения.

И така, може да се твърди, че с предложените в теста задачи се проверява нивото на подготовка на абитуриентите от средните учебни заведения по основните съдържателни линии на училищния курс по математика.

Нека да отбележим факта, че абсолютно същите задачи бяха предложени на кандидат-студентите и през 1994 г. от автора, което дава повод да се направят сравнения и изводи.

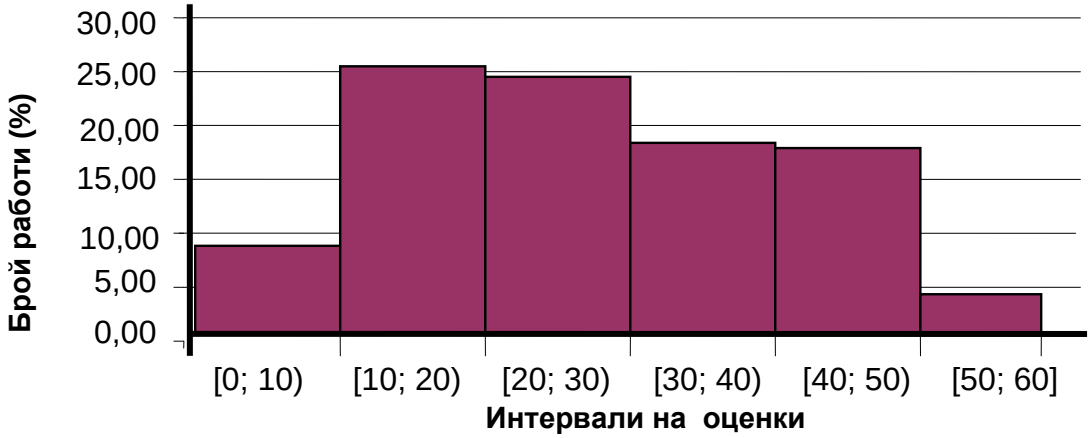
Получените от нас резултати от тестирането са представени във вид на диаграми и таблици (виж Диаграма 1, Диаграма 2 и Таблица 1).

Резултати от тестването **Диаграма 1**



Обработката на получените резултати показа, че средният успех е 26 точки

Резултати от тестването **Диаграма 2**



Резултати от тестирането **Таблица 1**

		1	2	3	4	5	6	7	8	9	10	Година
Задачи от първо ниво	**Брой правилни отговори (%)**	90,35	85,09	84,65	48,68	55,7	83,33	51,75	66,67	42,98	62,28	2005

Задачи от второ ниво	**Брой правилни отговори (%)**	82,46	56,14	52,63	37,72	31,58	28,95	32,02	42,54	20,18	42,11	2005
Задачи от трето ниво	**Брой правилни отговори (%)**	24,56	31,14	46,05	24,12	30,70	16,67	20,61	10,09	24,56	13,60	2005

По-детайлният анализ на резултатите от тестирането дава основания да се твърди, че:

а) успеваемостта по математика на абитуриентите от средните училища, както и качеството на тяхната математическа подготовка чувствително се е снизила.;

б) по всяка от предложените задачи от теста процентът на правилните решения през 2005 г. е по-малък, отколкото е бил например през 1994 г., при което често това намаление е почти двойно.

Тъжната и тревожна картина ни кара да се замислим над това, тези ли резултати се надявахме да получим в резултат от реформирането? Навярно, не такива. Дълбоко съм убеден, че реализирането в практиката на принципа за хуманитаризация на съдържанието на обучението и хуманизиране на учебния процес е отишло в Украина толкова далече, че пострада фундаменталната подготовка на абитуриентите в училище (подготовката не само по математика, но и по физика, химия и други дисциплини). Това се чувства по време на подготовката на специалистите за всички професии, особено тези, в които математиката се явява основа на тяхното професионално създаване. Считам, че още днес е необходимо, веднага, преди всичко:

1. На математиката да се върне статуса на водеща учебна дисциплина в училище. Достатъчно й беше да е "Пепеляшка". Всички ученици, задължително трябва да полагат матура по математика, както в 9, така и в 11 клас. Не трябва тези изпити да се подменят с външно тестиране на абитуриентите, то има друга цел. Уместно е да напомним, че в развитите западни страни завършващите средно училище се явяват на два задължителни държавни изпита – държавен език и математика. И в Украина трябва да бъде същото.

2. Да се въведе системата от преводни изпити[1] по математика (и не само по математика), както в основното, така и в средното училище. Нали подготовката за тези изпити е и обучение, и преговор, и закрепване, и систематизация и обобщение на изученото. Също така да се възстанови оценяването на резултатите от обучението на учениците за срока като стимул за повишаване на тяхната успеваемост. Без това е невъзможно да се повиши качеството на образованието.

3. Да се усъвършенства образователният стандарт по математика, по ясно да се опишат в него изискванията към математическата подготовка на учениците. Желателно е тези изисквания да бъдат изразени също във формата на задачи (така, както те се представяха в края на 80-те години в СССР във вид на "задължителни резултати от обучението".) С държавните нормативни документи да се осигури възможност за получаване от завършващия средно учебно заведение на диплома за средно образование само в този случай, когато резултатите от неговото обучение не са по-ниски от достатъчното ниво от изисквания на стандарта. Нали сега почти всички завършващи училище стават студенти във висшите

[1] Изпити за преминаване от един клас в друг. У нас с подобно съдържание се използва понятието "контролна работа за определяне на входно (изходно) ниво". Бел.пр.

учебни заведения с държавна или частна форма на обучение. Нима може да учи във висше учебно заведение човек, който няма средно образование?

4. Да се усъвършенстват, опирайки се на образователния стандарт, съществуващите и да се създават нови учебни програми по математика, учебници, учебни пособия, които биха съответствали на профила на обучение, осигурявайки възможности за осъществяване на диференциация по нива и профили.

5. Да се увеличи броят на часовете през седмицата за изучаване на математика в общообразователното, в частност в средното училище, за сметка на часовете от училищната компонента[2].

ЛИТЕРАТУРА

1. Математика. Программы для общеобразовательных учебных заведений. - К.: Учебная книга, 2003. – 302 с.
2. Книга для учителя математики: Справочно-методическое издание / Состав. Н.С. Прокопенко, Н.П.Щекань. – Харьков: ТОРСІНГ ПЛЮС, 2005. – 272 с.
3. Математика: Программа для общеобразовательных учебных заведений, 5–12 классы. – Киев, Ирпень: Перун, 2005. – 64 с.

[2] Резервните часове в учебния план, с които разполага всяко училище. Бел.пр.

ФОРМИРОВАНИЕ КОМПОНЕНТОВ ТВОРЧЕСКОГО МЫШЛЕНИЯ УЧАЩИХСЯ ПРИ ИЗУЧЕНИИ МАТЕМАТИКИ

ОЛЬГА С. ЧАШЕЧНИКОВА

THE FORMATION OF COMPONENTS OF CREATIVE THINKING IN THE STUDY OF MATHEMATICS

OLGA S. CHASHECHNIKOVA

The author proposes the possibility to use the study of mathematics to form components of creative thinking in non-special classes.

KEY WORDS: education at mathematics, creative thinking

Современный специалист рассматривается сегодня как субъект собственной профессиональной деятельности, как активная, свободная и ответственная за свои решения в проектировании, осуществлении, преобразовании собственной деятельности личность.

Еще более тридцати лет тому назад В.Н.Пушкин отмечал, что в профессиях, выдвигающих высокие требования к находчивости человека, уровень обученности играет значительно меньшую роль, чем способность быстро принимать ответственные решения в сложной оперативной ситуации [7]. Поэтому выпускники школ и высших учебных заведений должны не просто получить разносторонние глубокие и прочные знания, но и приобрести умения их применять творчески, нестандартно, постоянно пополнять систему знаний; находить оригинальные методы, способы, приемы решения задач и проблем; критически оценивать результаты собственной деятельности, усовершенствовать ее и себя в ней.

Не является новой идея о том, что система образования должна заботиться о развитии творческих качеств так же, как она традиционно заботится об интеллектуальных и профессионально-ориентированных умениях (R.Miller) [11]. Одной из приоритетных целей системы образования, построенной на гуманистических основах, должна быть цель развития творческого мышления обучающегося.

Однако анализ современного состояния обучения (в том числе, обучения математике) свидетельствует о недостаточной реализации этой идеи. Тенденция к профилизации обучения в школе, на наш взгляд, нередко сопровождается снижением внимания к возможностям, которые предоставляет для развития творческой личности учащегося, его творческого мышления изучение предметов, не входящих в состав “профильного ядра”.

Нерешенной еще на данном этапе является проблема развития творческого мышления учащихся классов нематематического профиля в процессе изучения математики. Чаще всего среди причин называются загруженность учащихся предметами, соответствующими профилю; невозможность за время, предоставляемое учебными планами на изучение математики в этих классах, использовать специально подобранный материал, направленный на развитие творческого мышления.

Цель данной статьи: продемонстрировать возможности использования программного материала для развития творческого мышления учащихся в процессе изучения математики.

Творческое мышление является одновременно и условием, и результатом творческой мыслительной деятельности. На наш взгляд, творческое мышление предполагает наличие как творческих, так и интеллектуальных способностей.

Математика как учебный предмет несомненно имеет огромное значение для развития творческого мышления учащихся вне зависимости от того, в классе какого профиля (математического или нет) они обучаются. Необходимо отметить, что в процессе учебно-познавательной деятельности по математике выделяются:

- стереотипы опыта, которые формируются, когда учитель (учебник) постоянно акцентирует внимание на определенных аспектах, происходит систематическая отработка определенного "набора действий", операций;

- "приобретения" учащихся, осуществляемые при обзорном сообщении некоторых фактов (внимание на них не концентрируется), когда опыт выполнения некоторых действий и операций является "попутным" (но не второстепенным) продуктом деятельности в процессе выполнения определенного задания.

При этом вовсе не обязательно привлекать материал вне программы для реализации цели развития творческого мышления учащихся. Тем более что, как показывает практика, в классах нематематического профиля это не всегда целесообразно. В процессе обучения математике при использовании программного материала нередко создаются ситуации, которые можно и необходимо эффективно использовать с целью развития творческого мышления учащихся.

Более реально осуществимым в классах нематематического профиля является путь, когда:

- нестандартно подается стандартный материал;

- нестандартно решаются стандартные, на первый взгляд, задания.

Целесообразно использовать системы упражнений, ярко демонстрирующие учащимся необходимость отказаться от привычных методов и способов, которые не срабатывают в определенных условиях, а также способствовать тому, чтобы учащиеся увидели возможность использования нестандартных подходов к их решению.

Рассматривая решение уравнений, неравенств и их систем в системе упражнений целесообразно выделить такие типы:

1. Задания, в которых нахождение области допустимых значений уравнения (неравенства, системы уравнений или неравенств) "сообщает" о нецелесообразности дальнейшего процесса их решения. Например:

а) $\sqrt{2-x}+\sqrt{x-9}=7$;

б) $\lg(x-8)>\lg(8-x)$.

Для осуществления эффективного решения этих заданий учащиеся должны увидеть, что достаточно найти область допустимых значений. Область допустимых значений уравнения 1а) – пустое множество; область допустимых значений неравенства 1б) – пустое множество.

2. Задания, в которых нахождение множества значений выражений A(x) и B(x) "информирует" об отсутствии корней уравнения A(x) = B(x) (неравенства $A(x)\le B(x)$, $A(x)\ge B(x)$ и т.д.).

Например:

а) $\sin x+\sin 2x+\sin 3x+\sin 4x+\sin 5x=6$;

б) $\sin 3x+8<\cos\frac{x}{2}$;

в) $5^x+10=-2\cdot 25^x$

Для решения уравнения 2а) достаточно определить, что для любого значения аргумента α $|\sin\alpha|\le 1$. Поэтому значение выражения в левой части при любом значении аргументов не превышает 5. Равенство невозможно.

Для решения неравенства 2б) учащиеся определяют, что значение выражения в левой части не меньше 7, а значение выражения в правой части при любом значении аргумента не превышает 1. Неравенство невозможно при любом значении аргументов.

На первый взгляд, уравнение 3в) наиболее эффективно решать, сводя к квадратному уравнению. Однако достаточно рассмотреть функции $f(x)=5^x+10$ и $g(x)=-2\cdot 25^x$. Множество значений первой функции $(10;+\infty)$, множество значений функции $g(x)=-2\cdot 25^x$ - $(-\infty;0)$. Равенство невозможно ни при каких значениях переменной х.

3. Задания, в которых графическое моделирование помогает выявить количество корней уравнения (системы уравнений) или их отсутствие. Например:

а) $1-\sqrt{x}=\sqrt{x+3}$;

б) $\frac{2}{|x|}\ge -x^2$;

в) $\begin{cases} x^2+y^2=9, \\ x^2-y=3 \end{cases}$.

Рассмотрим решение задания 3а). Построение графиков функций $y=\sqrt{x+3}$ и $y=-\sqrt{x}+1$ ярко иллюстрирует:

-функция $y=\sqrt{x+3}$ - монотонно возрастает; множество ее значений $[0;+\infty)$;

-функция $y=-\sqrt{x}+1$ монотонно убывает; множество ее значений $(-\infty;1]$.

При х=0 первая функция принимает значение $\sqrt{3}$, а вторая 1 ($\sqrt{3}>1$).

Графики функций $y=\sqrt{x+3}$ и $y=-\sqrt{x}+1$ не пересекаются. Учащимся несложно сделать вывод, что уравнение не имеет корней. Необходимо акцентировать внимание учащихся на том, что при выполнении заданий такого типа достаточно делать лишь эскизы графиков функций.

Решение задания 3б может быть быстрым и изящным именно при применении эскизов графиков функций $y=\frac{2}{|x|}$ и $y=-x^2$. В этом случае не требуется раскрытие знака модуля и соответствующие преобразования.

Для выяснения количества корней системы уравнений 3в достаточно в одной системе координат изобразить окружность – график уравнения $x^2+y^2=9$ и график функции $y=x^2-3$. Точка пересечения графиков $(0;-3)$ находится легко.

Точки пересечения графика функции $y=x^2-3$ с осью абсцисс (нули функции) таковы: $(\sqrt{3};0)$ и $(-\sqrt{3};0)$. Целесообразно сделать прикидку: $\sqrt{3}<3$ и $-\sqrt{3}>-3$. Вывод: система имеет три корня.

4. Задания, в которых уточнение области допустимых значений переменной необходимо на всех этапах преобразований.

Пример задания этого типа:

$\sqrt{8-x}+\sqrt{x-3}<\sqrt{x}$.

Если при нахождении области допустимых значений учитывается только знак подкоренного выражения ($x\in[3;8]$), это приводит к ошибочному решению задания. Необходимым является уточненение области допустимых значений на этапе перехода к неравенству $2\sqrt{8-x}\sqrt{x-3}<x-5$.

5. Задания, в которых множество решений совпадает с областью допустимых значений.

Пример. Решить неравенства:

а) $\sqrt{x-7}\ge\sqrt{7-x}$;

б) $\sqrt{x-9} \le \sqrt{x}$.

Решение таких неравенств целесообразно сопровождать графической иллюстрацией.

Данные задания можно предлагать учащимся как для индивидуального решения, так и для решения в группах. Работа над ними пробуждает заинтересованность учащихся, формирует и развивает их интеллектуальную активность, дает им опыт ведения аргументированной дискуссии в процесе “защиты” способа, который конкретный учащийся считает наиболее рациональным, развивает специфическое “математическое видение ”.

Результаты использования разработанной нами системы упражнений на уроках алгебры и начал анализа свидетельствует, что работа с ней не только способствует углублению знаний и повышению качества умений, успеваемости учащихся, но и делает более продуктивным процесс формирования у них способности отыскивать наиболее рациональные способы решения, развивает их математическую интуицию.

Приобретенный опыт поиска нестандартных путей мышления, повышение уровня интеллектуальной компетентности являются необходимыми при изучении не только математики, но и других предметов. Как следствие – продолжается процесс формирования и развития творческой личности учащегося.

На этапе изучения нового материала также можно удачно обыграть стандартный программный материал.

Рассмотрим на примере введения системы аксиом стереометрии. Учащимся можно предложить проанализировать систему аксиом, в которой в качестве аксиомы плоскости выступает формулировка относительно задания единственной плоскости (эта аксиома названа в учебнике [1] “аксиомой плоскости”):

двумя пересекающимися прямыми [3], [6] (обозначим эту аксиому C_3-1);

прямой и не принадлежащей ей точкой (обозначим C_3-2);

тремя точками, не лежащими на одной прямой [4] (обозначим C_3-3);

двумя параллельными прямыми (обозначим C_3-4).

Учащимся сообщают в доступной форме основные требования к системе аксиом. Одно из обязательных требований к системе аксиом – ни одна из аксиом не может быть следствием другой аксиомы этой же системы. Если одна из четырех вышеуказанных формулировок будет считаться формулировкой аксиомы, то любая из трех оставшихся – следствие из этой аксиомы, теорема. Поэтому учащимся можно предложить рассмотреть каждый из четырех случаев (не только тот, когда аксиомой считается C_3-1, а C_3-2, C_3-3, C_3-4 являются следствиями из нее, как в учебнике А.В.Погорелова, но и когда аксиомой считается C_3-2, а C_3-1, C_3-3, C_3-4 – следствия из аксиомы, и т.д.).

Таким образом создаются условия для развития творческого мышления учащихся:

- им предлагается нестандартно подойти к самостоятельному “конструированию” системы аксиом;

- предоставляется возможность самостоятельного поиска доказательства теорем с опорой на нетрадиционную формулировку аксиомы плоскости.

Кроме того, доказательство первых теорем стереометрии не только требует достаточного уровня сформированности абстрактного мышления, но и способствует эффективному его развитию, и, как следствие, развитию творческого мышления учащихся.

Необходимо предлагать учащимся задания, эффективность решения которых зависит от «ракурсов» рассмотрения условия. Существующие учебники содержат такие задания, необходимо только их найти и «обыграть».

Рассмотрим такое задание по теме «Объемы многогранников»: «Боковые ребра треугольной пирамиды взаимно перпендикулярны, каждое равно b. Найдите объем пирамиды» (№35 §21 [6]).

Нередко учащиеся начинают решение этой задачи по алгоритму, включающему этапы нахождения длин сторон основания пирамиды; нахождения длины высоты пирамиды. Такое решение часто бывает достаточно громоздким. Это – формальный, шаблонный подход (и «подсказывается» он чаще всего рисунком к задаче, тем ракурсом, который предложен в учебнике). Однако нет необходимости тратить много времени и усилий на решение вышеуказанной задачи, если «развернуть» пирамиду.

Допустим, пирамиду на предложенном в учебнике рисунке обозначим SABC, где SA, SB, SC – боковые ребра, ABC –основание пирамиды. Остроумным решением будет такое, которое основывается на рассмотрении, например, такого ракурса: SAB – основание пирамиды (равнобедренный прямоугольный треугольник с известными катетами), CS, CA, CB – боковые ребра, причем CS –высота пирамиды, длина которой известна.

Важно, чтобы учитель «спровоцировал» учащихся на выполнение рисунка, отличного от того, который представлен в учебнике. Это можно сделать опосредованно, демонстрируя соответствующую модель пирамиды.

Следующим этапом в решении этой задачи является предоставление учащимся возможности самостоятельно сопоставить оба предложенных подхода к решению, выбрать наиболее красивое из них, аргументировать это.

Однако подчеркнем, что ограничиваться этим нецелесообразно. Даже красивый, рациональный способ, вызвавший интерес в данный момент у учащихся, может оказаться недейственным в последствии, если учитель специально не акцентирует внимание учащихся на таком возможном нестандартном подходе к решению задач. Иначе этот найденный оригинальный способ, «попутно приобретенный» опыт, способствующий более эффективной работе, может оказаться кратковременным незакрепленным успехом учащихся. Для будущей действенности такого подхода полезно мотивировать учащихся к «внесению» этого способа в их «оперативную интеллектуальную базу», подчеркнув возможность использования его в будущем.

С этой целью можно предложить учащимся выполнить следующее задание: изобразить несколько фигур в пространстве, к элементам которых можно применить теорему о трех перпендикулярах. Таким образом решается несколько задач:

1. Происходит активное повторение ранее изученного материала при решении задачи, причем оно идет не от заданного условия к применению теоретического материала, а наоборот, что требует нестандартности хода мысли. Еще С.Л.Рубинштейн отмечал, что даже для проявления реминисценции (улучшения воспроизведения со временем) [8,337] наиболее эффективным является не буквальное воспроизведение, а свободное изложение смыслового содержания логического, а не иллюстративного характера.

2. Элементы фигур, которые можно переосмыслить в контексте использования теоремы о трех перпендикулярах (наклонная, перпендикуляр, проекция наклонной) могут на рисунках рассматриваться в непривычных ракурсах.

Причем эксперимент показывает, что часто учащиеся сначала изображают наклонную, перпендикуляр, проекцию наклонной, затем «дополняют» рисунок, а уже после выбирают удобный ракурс для рассмотрения самой фигуры. При решении же «готовой задачи» процесс идет в обратном направлении: фигуру рассматривают в таком ракурсе, чтобы удобнее было выделять наклонную, перпендикуляр, проекцию наклонной для применения теоремы о трех перпендикулярах.

3. Происходит переосмысление уже известного теоретического материала, причем в данном случае учащийся уже не пользуется «чужим» пониманием этого материала.

На данном этапе учащийся находится в иной ситуации, чем та, когда происходила систематизация актуализируемых сейчас знаний. Эта ситуация характеризуется тем, что в базе знаний и умений учащегося уже произошли определенные изменения (накоплены новые

знания, сформированы новые умения). И теперь учащийся имеет возможность рассматривать имеющуюся у него систему знаний и умений, соответствующих теме, в разных ракурсах, в том числе в тех, в которых он не мог бы ее рассматривать без наличия знаний и умений, полученных после изучения данного материала (в данном случае, после изучения теоремы о трех перпендикулярах). Психологи отмечают, что необходимо научиться быть открытым к содержанию и способным самостоятельно его оценивать, иначе можно утратить важную информацию (К.Андреас, С.Андреас) [2]. Подчеркнем – оценивать самостоятельно, а не только пользоваться оценками и суждениями других (учителя).

Творчество предполагает свободу от шаблонов, отсутствие боязни отличаться от стандартов, определенную смелость суждений. На практике учащиеся классов нематематического профиля нередко внутренне занижают оценку своих возможностей решать задания творческого характера. Поэтому не менее важным в процессе развития творческого мышления учащихся является приучение их не отбрасывать сразу же то, что пугает новизной, нестандартностью представления. Для этого эффективно использовать задания, суть которых остается стандартной, но скрывается за нестандартностью формы.

Например, учащихся классов нематематического профиля чаще всего на первом этапе «пугает» задание, связанное с необходимостью построить график:

$$y = \frac{x+\pi}{|x+\pi|} \cdot \left(\cos\left(x+|x|\right)+\cos\left(x-|x|\right)\right).$$

Эксперимент показывает, что при наличии свободы выбора из нескольких заданий, данное не выбирается значительной частью учащихся, даже при условии того, что за его выполнение можно получить большее количество баллов, чем за выполнение всех «стандартно представленных» вместе.

Учителю важно продемонстрировать решение этого задания учащимся, максимально привлекая их к процессу решения, для того, чтобы раскрепостить их, раскрыть перед ними их собственные возможности, создать ситуацию успеха, порождающую в дальнейшем стремление к творчеству.

Не является открытием то, что развитие творческого мышления учащихся в процессе изучения математики является наиболее эффективным именно тогда, когда оно не происходит стихийно, время от времени. Учитель не может полагаться только на то, что само по себе изучение математики как учебного предмета не может не способствовать формированию и развитию творческой личности учащегося. Конечно, возможности развития творческого мышления при изучении математики создаются самим содержанием и логикой этого учебного предмета, но этого, естественно, недостаточно. Как известно, школа воспитывает и развивает учащихся не только содержанием изучаемого материала, но и организацией и методами обучения (И.Т.Огородников) [5].

Действительно высокие результаты возможны, если развитие осуществляется целенаправленно, систематически, дифференцированно и индивидуализированно в духе сотрудничества, сотворчества учителя и ученика. Планирование, организацию, руководство этим процессом осуществляет учитель математики, который и сам должен быть творческой личностью.

Поэтому будущих учителей математики мы систематически знакомим с возможностями целенаправленного развития творческой личности учащегося в процессе обучения математике в школах разного типа, в классах разного профиля.

На занятиях по методике обучения математике мы демонстрируем, каким образом можно использовать стандартные на первый взгляд упражнения, содержащиеся в существующих учебниках и учебных пособиях по математике, для формирования и развития интеллектуальной компетентности учащегося.

При этом акцент делатся на том, чтобы научить студентов классифицировать эти задания, структуировать их систему соответственно заданной цели, учитывая профиль обучения, индивидуальные особенности учащихся конкретного класса. У будущих учителей математики формируется способность работать в школах разного типа, по разным программам, используя разные учебники и учебные пособия, и при этом осуществлять целенаправленный процесс развития творческой личности учащихся.

Дальнейшего изучения требует процесс преемственности развития творческого мышления в школе и высших учебных заведениях.

ЛИТЕРАТУРА

1. **Александров** А.Д., Вернер А.Л., Рыжик В.И. Геометрия для 10-11 кл.: Учеб. пособие для учащ. шк. и классов с углубленным изучением математики. – М.: Просвещение, 1992. – 464 с.
2. **Андреас** К., Андреас С. Измените свое мышление и воспользуйтесь результатами: Новейшие субмодальные вмешательства НЛП / Пер. с англ.- СПб: Ювента, 1994.- 238 с.
3. **Атанасян** Л.С., Бутузов В.Ф., Кадомцев С.Б., Киселева Л.С., Позняк Е.Г. Геометрия: Учеб. для 10-11 кл.сред.шк.- 2-е изд.- М.: Просвещение, 1993.- 207 с.
4. **Клопский** В.М., Скопец З.А., Ягодовский М.И. Геометрия: Учеб. пособие для 9-10 кл.сред.шк. – М.: Просвещение, 1982.- 256 с.
5. **Огородников** И.Т. Воспроизводящее и творческое овладение знаниями // Воспроизводящая и творческая деятельность учащихся в обучении. Сб.трудов. – М.: МГПИ, 1978. – С.3-16.
6. **Погорелов** А.В. Геометрия: Учеб. для 7-11 кл. сред.шк.- М.: Просвещение, 1990.- 384 с.
7. **Пушкин** В.Н. Психология и кибернетика.- М.: Педагогика, 1971.- 232 с.
8. **Рубинштейн** С.Л. Основы общей психологии: В 2т.-Т.1.-М.:Педагогика,1989.-488с.
9. **Чашечникова** О.С. Развитие интеллектуальной компетентности старшекласников в процессе изучения математики // Теория и методика обучения математике, физике, информатике. – Ч.I. – Кривой Рог, 2002. – С.402-406. (на укр. яз.)
10. **Чашечникова** О.С. Проблема направленности учебника геометрии на формирование и развитие рациональности мышления учащихся // Проблемы современного учебника: Сб.науч. трудов. – Вып.3. – Киев: Педагогічна думка, 2003. – С.94-96. (на укр. яз.)
11. **Miller** R. What Are School For? Holistic Education in American Culture. – Brandon, Vermont, USA, 1992.-P.153.

СИСТЕМА ИЗУЧЕНИЯ ИСТОРИИ МАТЕМАТИКИ

ВАЛЕНТИНА Г. БЕВЗ

A SYSTEM FOR STUDYING HISTORY OF MATHEMATICS

VALENTINA G. BEVZ

This article treats the history of mathematics as a school subject in the pedagogical university. It analyses the main goals of the course and its great importance to the teachers in mathematics. The author proposes the use of a complex approach to the teaching in history of mathematics and the use of computer based tests to evaluate students knowledge.

KEY WORDS: high education, history of mathematics

На современном этапе развития высшего педагогического образования история науки непосредственно влияет на педагогическую культуру учителя, существенно расширяя его общий и профессиональный кругозор, повышая уровень интеллекта и эрудицию, формируя современную систему ценностей и профессиональное поведение. Она имеет большое значение для формирования критического мышления, творческих способностей, научного мировоззрения, интереса к научным знаниям, а также для воспитания моральных качеств будущего учителя.

В педагогических университетах Украины все студенты изучают историю философии, историю психологии и историю педагогики как составляющие части соответственных учебных дисциплин. Кроме этого, учебные планы предусматривают изучение отдельным предметом истории профессиональной дисциплины. На физико-математических факультетах педагогических университетов изучаются история математики, история физики, история экономики и история информатики. То есть в педагогических университетах изучаются истории отдельных наук, а не история науки вообще как целостной системы человеческих знаний.

Остановимся подробнее на проблеме изучения истории математики в педагогическом университете на современном этапе, а именно: определим цель курса, раскроем его значение для учителя математики, а также рассмотрим методические особенности построения курса и его реализация в учебном процессе.

Цель изучения курса “История математики” в педагогическом университете имеет два взаимосвязанных аспекта – общенаучный и профессиональный. Общенаучная цель изучения курса “История математики” состоит в том, чтобы показать общую картину возникновения и развития математических понятий, методов и теорий; выяснить характер и особенности развития математики у различных народов в определенные исторические периоды; показать вклад, внесенный в математику большими учеными прошлого; продемонстрировать многогранные связи математики: с практическими потребностями и деятельностью людей, с развитием других наук, а также влияние экономической, социальной и идеологической структуры общества на характер развития математики; раскрыть историческую обусловленность логической структуры современной математики, диалектику ее развития, соотношение частей математики и ее перспективы.

Цель изучения курса “История математики” в рамках профессиональной подготовки учителя математики состоит в том, чтобы дать будущим учителям исторические знания, необходимые им для правильного решения методологических вопросов, которые возникают в процессе преподавания математики; раскрыть глубокие психолого-методические вопросы

взаимосвязи истории науки и практики школьного обучения; показать место, значение и возможности исторического материала в школьном курсе математики современной школы.

Значение истории математики для современного учителя определяется содержанием и методологическим направлением самого курса и требованиями времени. Историко-математическая подготовка студентов оказывает воздействие на формирование научного мировоззрения, создает правильное представление о математике как о целостной, постоянно развивающейся науке, что дает возможность проследить, как возникали основные математические понятия, идеи и методы, как формировались отдельные математические теории в тот или иной исторический период. Важную роль курс истории математики играет в решении проблемы гуманизации и гуманитаризации образования, так как фактически она является гуманитарной составляющей математического знания. С ее помощью, используя биографии творцов математики, их методы работы, достижения и промахи, достаточно просто и убедительно удается одухотворить и очеловечить математику, повлиять на формирование характера учеников, их идеалов и стремлений.

Будущий учитель математики должен не только знать классические математические дисциплины, но и понимать структуру математики в целом, иметь представление о современных областях математики, понимать ее методологические проблемы, знать связи математики с другими науками и ее разнообразные применения. Все эти вопросы в систематизированном виде раскрываются в процессе изучения курса „История математики”.

Историю математики как науку нужно отличать от учебного курса истории математики. Содержание последнего формируется из большого объема исторических сведений, накопленных и систематизированных наукой. При этом основное внимание уделяется материалу, который будет оказывать содействие на формирование математической культуры и профессиональной компетентности будущего учителя математики.

Содержание курса истории математики на физико-математических факультетах педагогических университетов также отличается от содержания этого курса в классических университетах. Будущий учитель математики должен получить такие же знание общих вопросов истории математики, как и другие математики, но кроме этого овладеть определенным кругом узких вопросов по истории развития математики, которые непосредственно используются в процессе преподавания математики в школе. Кроме этого, желательно вооружить будущего учителя рациональными приемами использования полученных знаний в профессиональной деятельности: на уроках, во внеклассной работе, в учебно-воспитательном процессе вообще.

Систематический курс “История математики” в педагогических университетах изучается преимущественно на последних курсах. Это связано с тем, что изучение истории математики как науки возможно лишь после овладения студентами основных математических дисциплин: алгебры, геометрии, математического анализа, теории вероятности и тому подобное. Как раз курсом истории математики завершается общая математическая подготовка студентов на физико-математическом факультете. На его изучение отводится 54 часа (18 - самостоятельная работа студентов, 24 - лекции, 12 - практические) в 10 семестре. Такое место курса истории математики в учебном процессе не оказывает содействие достижению поставленной цели, в особенности ее профессионального аспекта, а именно: студенты не имеют возможности в полной мере использовать знания по истории математики в процессе написания курсовых работ по различным математическим дисциплинам, на лабораторных занятиях по методике математики, во время прохождения педагогической практики и тому подобное. Значительную часть исторического материала студентам приходится обрабатывать самостоятельно.

В Национальном педагогическом университете имени М. П. Драгоманова разработан комплексный подход к изучению истории математики. Процесс усвоения знаний по этой

дисциплине происходит в той или иной мере на лекциях и семинарских занятиях, во время самостоятельной работы с литературой, написания рефератов, выполнения индивидуальных заданий и сдачи зачета.

Пропедевтическое изучение истории математики начинается на первом курсе. С этой целью для студентов математических специальностей вводится факультативный курс “Математика как наука и учебный предмет”, в котором, кроме других, рассматриваются вопросы: что такое математика; основные этапы развития математики; современная математика и ее архитектура; математика в системе других наук; роль математики в развитии общества; эволюция математического языка.

Отдельные историко-математические сведения студенты получают на лекциях и практических занятиях по элементарной математике, высшей математике, методике преподавания математики, физике, информатике и т.д.

На четвертом курсе (7 семестр) в рамках математического кружка рассматриваются вопросы, связанные с использованием элементов истории математики в школе. Студентам раскрываются общие вопросы использования историзмов в школьном курсе математики, подается содержание исторического материала, который можно использовать в процессе обучения математике в 5-6 классах, алгебры и геометрии в 7-9 классах, рассматриваются методические особенности использования такого материала. Студенты имеют возможность выяснить цель и место использования элементов истории математики на уроках и во внеклассной работе, получить рекомендации относительно отбора тем и конкретного исторического материала, ознакомиться с основными приемами его представления, рассмотреть исторические задачи и особенности их решения.

Насыщенность курса фактическим материалом (он охватывает развитие математики с древнейших времен и до XX столетия включительно) создают широкие возможности для его изложения. Существует несколько способов построения истории математики как учебного предмета, а именно: историко-хронологический, предметно-модульный, концептуально-логический (история математики как история идей), доминантный, персоналистический и комбинированные.

Традиционно лекционный курс истории математики строят на хронологической основе: 1) зарождение математики; 2) элементарная математика; 3) математика переменных величин; 4) современная математика. Основное внимание здесь уделяется фактам, гипотезам, теориям, законам, методологии, а также определению роли научных работников, в частности украинских, в развитии математики. Представление и полнота изложения отдельных тем определяется возможностями, вкусами и желаниями конкретного лектора.

Современные требования к учебному процессу в университете предусматривают комплексный подход к изучению истории математики. Процесс усвоения историко-математических знаний происходит на лекциях и семинарских занятиях, во время самостоятельной работы над соответствующей литературой, подготовки рефератов, решения исторических задач.

Подготовленная нами программа [1] предусматривает комбинированный подход к построению учебного курса “История математики”. Лекционный курс, за исключением первой и последней лекций, строится на хронологической основе. Это объясняется тем, что история математики есть не только математическая, но и историческая наука. Для хронологического способа определяющим есть время появления той или иной идеи, понятия или метода. Он позволяет проследить историю развития всей математики в целом, показать ее связь с историей культуры и общества на различных этапах их развития. При таком подходе хорошо видны взаимосвязь и взаимовлияние математики и других наук.

Выходят за пределы хронологического изложения первая и последняя лекции учебного курса. На первой (вступительной) лекции рассматриваются вопросы методологического

характера: предмет математики и истории математики; движущие силы развития математики; значение историко-математических знаний для учителя математики; основные этапы развития математики, а на последней – развитие математики в Украине.

Значительные сложности возникают в процессе построения практического курса. Последовательность и тематика семинарских занятий подбирается таким образом, чтобы они постоянно находились в тесной связи с лекционным курсом и в то же время существовала возможность рассмотреть на семинарских занятиях материал, который не изучается на лекциях. Наш опыт работы показывает, что с этой целью семинарские занатия лучше строить не на хронологической основе, а в соответствии с основніми разделам математики. А именно: история арифметики; история алгебры; история геометрии; история математического анализа; история развития отдельных разделов математики.

Для примера приведем план семинарского занятия по теме “История развития отдельных разделов математики”:

1. История развития комбинаторики.
2. Возникновение и развитие теории вероятностей.
3. Три источника векторного исчисления.
4. Основные этапы развития топологии.
5. История развития теории фракталов.
6. Исторические задачи.

На каждом занятии, кроме традиционного освещения вопросов по теме, практикуем отображать изученные сведения на материал школьного курса математики и решать соответствующие исторические задачи. Такой подход дает возможность будущему учителю математики на конкретных примерах увидеть место и роль исторического материала в учебном процессе, получить методические рекомендации относительно его использования и убедиться в необходимости исторических экскурсов во время изучения математики в школе.

Особое место на семинарских занятиях по истории математики должны занять исторические математические задачи. Решать их желательно несколькими способами: как их решали в эпоху создания и как их можно решать современными способами. Желательно чтобы в конце изучения курса истории математики у студентов сформировалась система исторических математических задач, построенная в соответствии с темами школьного курса математики.

Последнее семинарское занятие полностью посвящаем использованию элементов истории математики в школьной практике. В процессе подготовки к этому занятию студенты должны ознакомиться с историей проблемы, а для этого изучить рекомендованную историческую и методическую литературу, сделать анализ школьных программ, учебников и учебных пособий; разработать фрагменты уроков или внеклассных мероприятий с использования исторических сведений.

Отдельной проблемой обучения истории математики есть организация и проведение контроля знаний студентов. Как способ диагностики знаний, которые получили студенты после изучения курса “История математики”, мы предлагаем использовать компьютер. С этой целью нами была разработана тестовая система [2], состоящая из трех разделов:

- творцы математики;
- высказывание о математике и математиках;
- математическая мозаика.

Компьютер в системе произвольных чисел предлагает каждому студенту 10 вопросов из базы каждого раздела, а студент должен выбрать правильный ответ из четырех предложенных. По результатам тестирования компьютер оценивает знания студентов по 5-балльной шкале.

Организованная таким образом историко-математическая и историко-методическая подготовка студентов на физико-математическом факультете Национального педагогического университета им. М. П. Драгоманова оказывает реальное содействие в подготовке учителей высокого уровня образованности и профессиональной компетенции.

ЛИТЕРАТУРА

1. **Бевз** В.Г., Сазонова О.П. Програма з історії математики.// Програми з методики навчання математики, елементарної математики та історії математики. – К.: НПУ ім. М.Драгоманова, 2001. – с. 26–34.
2. **Бевз** В. Г. Історія математики: Тестові завдання для контролю знань: Навч.-метод. посібник у 2-х частинах. – К.: НПУ імені М. П, Драгоманова, 2004. – Методичні вказівки. Ч. II. – 18 с.

РАВНИННИ КРИВИ С ПОЛОЖИТЕЛНА ШЕЙП КРИВИНА

ГЕОРГИ ХР. ГЕОРГИЕВ, РАДОСТИНА П. ЕНЧЕВА

PLANE CURVES WITH A POSITIVE SHAPE CURVATURE

GEORGI HR. GEORGIEV, RADOSTINA P. ENCHEVA

The shape curvature of a plane curve is an invariant under the group of direct similarities. In this paper, we investigate parts of conic sections with a positive shape curvature. Moreover, they are found not trivial examples of plane curves that have positive shape curvatures for any point of curves.

KEY WORDS: shape curvature, similarity group, conic section, spirals, evolvent.

1. ВЪВЕДЕНИЕ

Дейвид Кендъл първи дефинира в [3] така нареченото шейп пространство като фактор пространство относно групата на подобностите. За изучаване на тези пространства се използват шейп инварианти, които остават постоянни относно всяко преобразуване от разглежданата група. Такава инварианта, определяща равнинна крива на Френе с точност до подобност, запазваща ориентацията, е въведената в [1] шейп кривина. Ще дадем дефиницията на тази функция.

Нека $c: I \to \mathbf{R}^2$, $c = c(\sigma)$, $\sigma \in I$ е равнинна крива на Френе, параметризирана спрямо сферичен естествен параметър σ, който е естественият параметър на сферичната допирателна индикатриса на кривата. Функцията

$$\tilde{K}(\sigma) = -\frac{dK}{d\sigma}.\frac{1}{K}, \qquad (0.1)$$

където K е равнинната кривина на c, е инварианта относно групата на подобностите, запазващи ориентацията. Освен това имаме, че

$$\tilde{K}(\sigma) = \frac{<\frac{dc}{d\sigma}, \frac{d^2c}{d\sigma^2}>}{<\frac{dc}{d\sigma}, \frac{dc}{d\sigma}>}.$$

Доказаната в [1] Теорема 2, е аналог на фундаменталната теорема в локалната теория на равнинните криви. Тя ни дава основание да наречем определената от (1.1) функция $\tilde{K}$, шейп кривина. Тази функция съответства на целия клас от подобни и еднакво ориентирани равнинни криви на Френе. Следващите формули изразяват шейп кривината на крива чрез естествения параметър s на кривата както и чрез произволен параметър t.

$$\tilde{K}(s) = \frac{d}{ds}\left(\frac{1}{K(s)}\right), \qquad (0.2)$$

$$\tilde{K}(t) = -\frac{dK}{dt}.\frac{1}{K^2\left\|\frac{dc}{dt}\right\|} = \frac{3<\frac{d^2c}{dt^2}, \frac{dc}{dt}>.<\frac{d^2c}{dt^2}, J\frac{dc}{dt}> - <\frac{dc}{dt}, \frac{dc}{dt}>.<\frac{d^3c}{dt^3}, J\frac{dc}{dt}>}{\left(<\frac{d^2c}{dt^2}, J\frac{dc}{dt}>\right)^2}, \qquad (0.3)$$

където J е комплексната структура в $\mathbb{R}^2$, а $<\cdot,\cdot>$ е стандартната метрика в $\mathbb{R}^2$.

Единствените криви, за които шейп кривината е постоянна във всяка точка от кривата, са окръжностите и логаритмичните спирали. При окръжностите тази функция е нула, а при логаритмичните спирали, положителна или отрицателна във всяка точка константа. В част 2 са

изследвани части от конични сечения, във всяка точка от които шейп кривината е положителна. Примери на равнинни криви, които имат положителна шейп кривина във всяка точка, са разгледани в част 3.

2. КОНИЧНИ СЕЧЕНИЯ

Известните от дълбока древност конични сечения са елипса, хипербола и парабола. Оказва се, че шейп кривината на тези равнинни криви не запазва знака си в различните им точки и се анулира във върховете. В следващите твърдения ще определим частите от коничните сечения, в които шейп кривината е положителна.

Твърдение 1 *Нека* $c: I \to \mathbf{R}^2$ *е елипса. Тогава*

a) Шейп кривината на c *е равна на нула единствено във върховете на елипсата.*

b) Шейп кривината на c *е положителна върху едната двойка дъги от елипсата, ограничени от осите и симетрични спрямо центъра ѝ.*

Доказателство. Ако $c(t) = (a\cos t, b\sin t), \quad t \in [0, 2\pi]$ е едно параметрично представяне на елипса с дължини на полуосите *a, b*, където *a>b*, то от (1.3) получаваме

$$\tilde{K}(t) = \frac{3(a^2 - b^2)\cos t.\sin t}{ab}.$$

Оттук, твърденията в a) и b) следват от свойствата на тригонометричните функции.

Твърдение 2 *Нека* $c: I \to \mathbf{R}^2$ *е хипербола. Тогава*

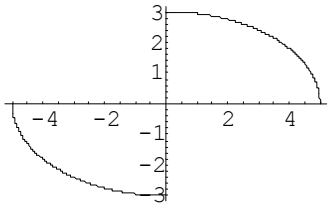
a) Шейп кривината на c *е равна на нула единствено във върховете на хиперболата.*

b) Шейп кривината на c *е положителна върху едната двойка клонове на хиперболата, ограничени от реалната ѝ ос и едната асимптота и симетрични спрямо центъра ѝ.*

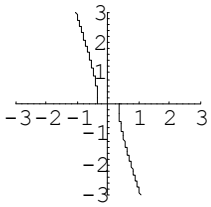
Доказателство. Ако $c(t) = (\pm a\cosh t, b\sinh t)$ е едно параметрично представяне на хипербола с дължини на полуосите *a, b*, то от (1.3) получаваме

$$\tilde{K}(t) = \mp\frac{3(a^2 + b^2)\cosh t.\sinh t}{ab}.$$

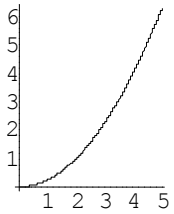
Оттук, твърденията в a) и b) следват от свойствата на хиперболичните функции.



Фиг. 1



Фиг. 2



Фиг. 3

Твърдение 3 *Нека* $c: I \to \mathbf{R}^2$ *е парабола. Тогава*

a) Шейп кривината на c *е равна на нула единствено във върха на параболата.*

b) Шейп кривината на c *е положителна върху едния клон на параболата.*

Доказателство. Ако $c(t) = (t, t^2/2p)$ е едно параметрично представяне на парабола с параметър *p>0*, то от (1.3) получаваме $\tilde{K}(t) = 3pt$. Оттук, твърденията в a) и b) са очевидни.

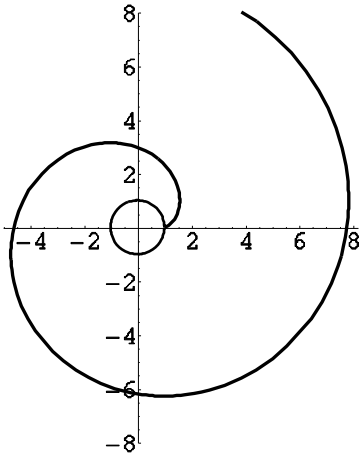
Доказаните твърдения показват, че за всяко от коничните сечения може да се намери интервал, в който шейп кривината да приема само положителни стойности. Фигурите 1, 2 и 3 илюстрират части от елипса, хипербола и парабола, съответно, където шейп кривината е положителна.

3. ЕВОЛВЕНТА НА ОКРЪЖНОСТ И СПИРАЛИ В РАВНИНАТА

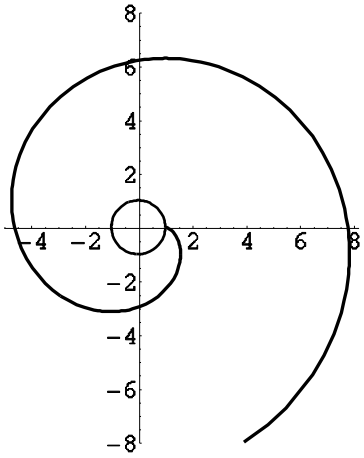
В тази част ще разгледаме равнинни криви, чиято шейп кривина има постоянен знак във всяка точка от кривата. Тривиален пример за равнинни криви с това свойство са логаритмичните спирали. Оказва се, че някои класове от спирали в равнината също се характеризират с тази особеност.

Нека $c = c(s), s \in I$ е правилна крива от клас C^3, параметризирана спрямо естествен параметър. Еволвента на c е кривата $\tilde{c}$ с векторно-параметрично уравнение $\tilde{c}(s) = c(s) + (C - s).T$, където T е единичният допирателен вектор в съответната точка на c, а C=const. Известно е, че еволвентите на окръжност с радиус a са равнинни криви с естествено уравнение $K(s) = \pm(2as)^{-1/2}$ (виж Фиг. 4 и Фиг. 5). Оттук прилагайки (1.2) получаваме шейп кривината на еволвента на окръжност, а именно $\tilde{K}(s) = \pm\dfrac{a}{\sqrt{2as}}$. Така е установена верността на твърдението

Твърдение 4 *Шейп кривината във всяка точка на еволвента на окръжност е или само положителна или само отрицателна.*



Фиг. 4



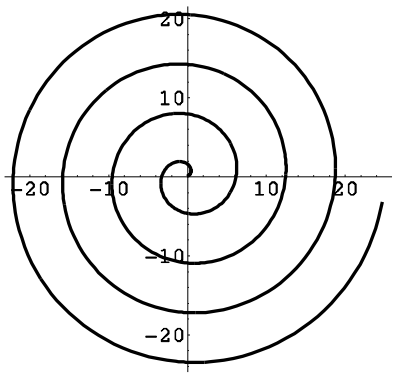
Фиг. 5

Нека $c(t) = (at^n \cos t, at^n \sin t)$, където $|n| > 1$, е параметрично представяне на обобщена спирала, параметризирана спрямо полярния ъгъл t (виж [2]). От (1.3) намираме, че шейп кривината на c е функцията

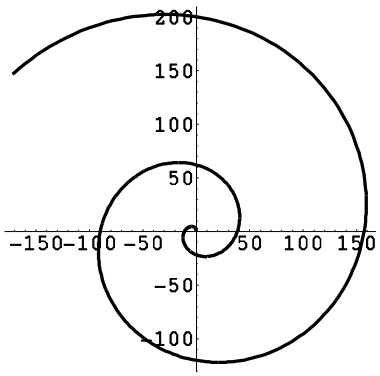
$\tilde{K}(t) = \dfrac{n(n^2(-1+n^2)+2(1+n^2)t^2+t^4)}{t(n+n^2+t^2)^2}$. Оттук следва твърдението

Твърдение 5 *Шейп кривината във всяка точка на обобщена спирала е или само положителна или само отрицателна.*

На следващите фигури са изобразени спирали, в случаите $n = 1$ (Архимедова спирала, Фиг. 6) и $n = 2$ (Галилеева спирала, Фиг. 7), с положителна във всяка точка шейп кривина.



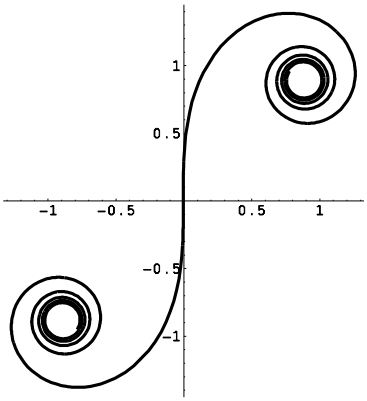
Фиг. 6



Фиг. 7

Накрая ще разгледаме един клас от спирали, за които шейп кривината приема само положителни стойности във всяка точка от кривата. Нека

$c(t) = \left(a\int_0^t \cos\left(\frac{u^{n+1}}{n+1}\right)du, a\int_0^t \sin\left(\frac{u^{n+1}}{n+1}\right)du \right)$, където $n > 0$, е параметрично представяне на обобщена спирала на Корну (виж Фиг. 8 в случая $n = 1$).



Фиг. 8

Твърдение 6 *Шейп кривината на обобщена спирала на Корну, за която* n *е нечетно число, е положителна във всяка точка от кривата.*

Доказателство. Прилагайки (1.3) и използвайки представянето по-горе имаме, че шейп кривината на обобщена спирала на Корну е функцията $\tilde{K}(t) = \frac{n}{t^{n+1}}$. □

ЛИТЕРАТУРА

1. **Encheva** R., G. Georgiev, *Curves on the shape sphere, Result. Math*., 44 (2003), 279-288.

2. **Gray** A., *Modern differential geometry of curves and surfaces*, CRC Press, Boca Raton, 1993.

3. **Kendall** D., *Shape manifolds, procrustean metric, and complex projective spaces*, *Bull. London Math. Soc.*, 16 (1984), 81-121.

ПРОСТИ КВАЗИТЪНКИ ГРУПИ ОТ НОРМАЛЕН ЛИЕВ ТИП

НИКОЛА Т. ПЕТРОВ

SIMPLE QUASI THIN GROUPS OF NORMAL LIE TYPE

NICOLA T. PETROV

It is checked what of the known finite simple groups are quasi thin (Theorem below). The proof is contained in this and the following two papers of the author.

KEY WORDS: finite simples group, quasi thin group

Ако M е крайна група, а p е просто число, p-ранг на групата M се нарича максималният между ранговете на абеловите й p-групи; ще го бележим с $m_p(M)$. Ако M е подгрупа на групата G и $M = N_G(E)$ за някоя 2-подгрупа $E \neq \{e\}$, M се нарича 2-*локална* група. Ако простото число p е фиксирано, с $m_{2,p}(G)$ ще бележим максимума на p-ранговете $m_p(M)$, когато M пробягва всички 2-локални подгрупи на групата G. Означаваме

$$e(G) = \max_{p>2} m_{2,p}(G),$$

където максимумът е по всички нечетни прости делители на реда $|G|$ на групата. Ако $e(G) \leq 1$ групата се нарича *тънка*, а при $e(G) \leq 2$ - *квазитънка*. С други думи G е квазитънка точно тогава, когато никоя от 2-локалните й подгрупи не съдържа елементарна абелова подгрупа от ред p^3 за някое нечетно просто число p.

През февруари 1981 г. беше анонсирано, че класификацията на простите крайни групи е завършена, но с уговорката, че предстои публикуването на класификацията на квазитънките групи. Класификацията на простите квазитънки групи е съществен раздел от доказателството на общата класификационна теорема за простите групи. До днес (октомври 2001 г.)[3] такава публикация не ни е известна (вж. и статията на R. Solomon [5]). В тази връзка изглежда не е лишено от смисъл да се провери акуратно кои от известните прости крайни групи са квазитънки. Заради относително големия обем на проверката ще я изложим в няколко отделни статии. Ще отбележим, че простите тънки групи бяха класифицирани от Aschbacher [1].

Главната ни цел е да докажем следната

Теорема. *Между известните неабелови прости крайни групи квазитънки са само следните:*

а) *групи на Шевалие над поле с четна характеристика:* $SL_2(q)$ *при* $q > 2$, $L_3(q)$, $SL_4(2)$, $SL_5(2)$, $U_3(q)$ *при* $q > 2$, $U_4(q), U_5(4)$, $Sp_4(q)'$, $Sp_6(2)$, $Sz(q)$ *за* $q = 2^{2l+1}$, $l \geq 1$, $G_2(q)'$, ${}^3D_4(q)$, ${}^2F_4(q)'$ *за* $q = 2^{2l+1}$ (*навсякъде* q *е степен на* 2);

б) *групи на Шевалие над поле с нечетна характеристика* p: $L_2(p^l)$, $p^l > 3$, $L_3(p)$, $L_3(p^2)$, $U_3(p)$, $U_3(p^2)$, $G_2(p)$, $PSp_4(p)$, $L_4(p)$ *за всяко просто число на Ферма;* $U_4(p)$ *за всяко просто число на Мерсен;*

в) *алтернативни групи:* $A_5 \cong L_2(4)$, $A_6 \cong L_2(9)$, A_7, $A_8 \cong L_4(2)$, A_9;

[3] Работата е докладвана на Юбилейната научна сесия по случай 30 годишнината на Шуменския университет "Епископ Константин Преславски", октомври 2001 год.

г) *спорадични групи:* M_{11}, M_{12}, M_{22}, M_{23}, M_{24}, J_1, J_2, J_3, J_4, HS, He, Ru, Mc.

Означения. Повечето от означенията, които ще използваме, са или стандартни, или класически, и няма да правим специални уговорки. Буквата p навсякъде означава просто число; $U_n(q)$ (или $PSU_n(q)$) е проективната специална унитарна група над поле с q^2 елемента; $L_n(q) = PSL_n(q)$; $S_p(L)$ или просто S_p е (някоя) силова p-подгрупа на групата L; $M_n(q)$ е пръстенът на квадратните матрици от ред n над поле с q елемента. Символът $diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ означава диагонална матрица с елементи $\lambda_1, \lambda_2, \ldots, \lambda_n$ по диагонала.

Доказателството на теоремата ще изложим във вид на серия от леми. За да избегнем повторенията се уславяме, че навсякъде A ще означава елементарна абелова група от ред p^3, E - елементарна абелова 2-група, като A и E са подгрупи на групата G и $A \subset N_G(E)$. Очевидно винаги можем да считаме, че ако A нормализира 2-група, то A нормализира и елементарна абелова 2-група.

Лема 1. *Нека A действа неприводимо на E. Тогава съществува подгрупа $A_0 \subset A$, $|A_0| = p^2$ и $A_0 \subset C_G(E)$.*

Доказателство. Ако абеловата p-група действа точно на векторното пространство E над поле с два елемента , то тя би била циклична, тъй като представянето е неприводимо и $(p,2) = 1$. Следователно $A/C_A(E)$ е циклична, т.е. $C_A(E)$ има ред поне p^2 и съдържа подгрупа A_0 с исканото свойство.

Лема 2. *В серията групи $L_n(q)$ при четно q квазитънки са само групите $SL_2(q)$, $L_3(q)$, $SL_4(2)$, $SL_5(2)$.*

Доказателство. В групата $SL_2(q)$ при четно q всички силови подгрупи от нечетен ред са циклични, следователно тя е тънка. Групите $L_4(2)$ и $L_5(2)$ са квазитънки, защото

$$|L_4(2)| = 2^6 \cdot 3^2 \cdot 5 \cdot 7, \quad |L_5(2)| = 2^{10} \cdot 3^2 \cdot 5 \cdot 7 \cdot 31$$

и, очевидно, силовите им подгрупи от нечетен ред имат ранг най-много 2. По-нататък имаме

$$|L_3(q)| = d^{-1} q^3 (q-1)^2 (q+1)(q^2+q+1), \quad d = (3, q-1).$$

Ако 3 не дели $q-1$, числата $q-1$, $q+1$, q^2+q+1 са две по две взаимно прости, а $L_3(q) = SL_3(q)$. Групата $SL_3(q)$ съдържа циклични подгрупи от редове $q+1$ и q^2+q+1, а също и директно произведение на две циклични от редове $q-1$. Следователно силовите подгрупи от нечетен ред имат ранг най-много 2, т. е. $e(L_3(q)) \le 2$. Ако 3 дели $q-1$, към горното разсъждение трябва да се добави разглеждане на силовата 3-подгрупа, защото сега $(q-1, q+1) = 1$, но $(q-1, q^2+q+1) = 3$. В $SL_3(q)$ силовата 3-подгрупа се поражда от диагонални матрици и от пермутационна матрица T, за която $T^3 = E$ (единичната матрица). Тя е неабелова и рангът й е равен на 2. При хомоморфизма $SL_3(q) \to L_3(q)$ рангът на силовата 3-подгрупа не може да се увеличи.

Нека $n > 3$ и $q = 2^l$. Групата $SL_n(q)$ съдържа директното произведение на $n-1$ циклични групи от редове $q-1$, което нормализира силовата 2-подгрупа (диагоналните матрици нормализират групата на горните триъгълни унипотентни матрици). Тъй като $SL_4(q) = L_4(q)$ при $n \ge 4$ се съдържа в $L_n(q)$, то при $q > 2$ веднага получаваме група A, която нормализира силовата 2-подгрупа.

Нека $q = 2$. Сега $L_n(2) = SL_n(2)$. Можем да предполагаме $n \ge 6$, защото вече разгледахме останалите случаи. Означаваме:

$$\lambda = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \qquad E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Очевидно $\lambda^3 = E$. Групата

$$A = \left\{ \begin{pmatrix} \lambda^i & 0 & 0 \\ 0 & \lambda^j & 0 \\ 0 & 0 & \lambda^k \end{pmatrix} | 1 \le i, j, k \le 3 \right\}$$

е елементарна абелова подгрупа (на $L_6(2)$) от ред 3^3 и нормализира 2-групата

$$L = \left\{ \begin{pmatrix} E & 0 & a \\ 0 & E & 0 \\ 0 & 0 & E \end{pmatrix} | a \in M_2(2) \right\}$$

от ред 16. Следователно $e(SL_6(2)) \ge 3$. Тъй като $n \ge 6$, то имаме $SL_6(2)$ $\subset SL_n(2) = L_n(2)$ и $e(SL_n(2)) \ge 3$.

Лема 3. *При нечетно q, $q = p^l$, групата $L_n(q)$ не е квазитънка с евентуално изключение на $L_2(p^l)$, $L_3(p)$, $L_3(p^2)$, $L_4(p)$. Ако $L_4(p)$ е квазитънка, то p е просто число на Ферма.*

Доказателство. Нека $n \ge 3$ и $l \ge 3$. Тогава групата $L_n(q)$ съдържа подгрупа $P_1 \cdot < t_1 >$, където $t_1 = diag(-1,-1,1,\ldots,1)$,

$$P_1 = \left\{ \begin{pmatrix} 1 & \lambda & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix} | \lambda \in GF(q) \right\}.$$

Тъй като групата P_1 е елементарна абелова от ред q с ранг $l \ge 3$ и централизира инволюцията t_1, то групата $L_n(q)$ при $n \ge 3$ и $l \ge 3$ не е квазитънка. В частност получаваме, че ако при $n = 3$ тя е квазитънка, то $l \le 2$.

Нека $n \ge 5$. Сега групата $L_n(q)$ съдържа подгрупа $P_2 \cdot < t_2 >$, $t_2 = diag(-1,-1,1,-1,-1)$,

$$P_2 = \left\{ \begin{pmatrix} 1 & \lambda & 0 & 0 & \nu \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & \mu \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} | \lambda, \mu, \nu \in GF(q) \right\}.$$

Групата P_2 е елементарна абелова от ред q^3, рангът й е $3l \ge 3$ и централизира инволюцията t_2. Следователно при $n \ge 5$ групата $L_n(q)$ не е квазитънка.

Нека $n = 4$. Ако $l \ge 2$, ще повторим горното разсъждение с инволюцията $t_3 = diag(-1,-1,1,1)$ и подгрупата

$$P_3 = \left\{ \begin{pmatrix} 1 & \lambda & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \mu \\ 0 & 0 & 0 & 1 \end{pmatrix} | \lambda, \mu \in GF(q) \right\}$$

и ще се убедим, че групата $L_4(q)$ при $l \ge 2$ не квазитънка.

Нека $n=4$ и $q=p$ не е просто число на Ферма. Разглеждаме групата на диагоналните матрици в $SL_4(p)$ и я факторизираме по подгрупата на скаларните матрици в $SL_4(p)$. Нека H е получената група. Редът на H е $d^{-1}(p-1)^3$, където $d=(4,p-1)$. Тъй като $p-1$ не е степен на 2, подгрупата $H\subset L_4(p)$ съдържа абелова подгрупа от нечетен ред и ранг 3, която централизира инволюция от H. Следователно $L_4(p)$ не е квазитънка.

Лема 4. *Нека p е произволно нечетно просто число. Тогава групите $L_2(p^l), L_3(p), L_3(p^2)$ са квазитънки.*

Доказателство. Редът на групата $L_2(q)$ е $\frac{1}{2}q(q^2-1)$, $q=p^l$. Тя съдържа циклични подгрупи от редове $\frac{1}{2}(q\pm1)$. Тъй като $(q-1,q+1)=2$, то всички силови r-подгрупи, $r>2$, са или циклични, или елементарни абелови p-групи. Но в $L_2(q)$ централизаторът на всеки p-елемент е p-група, което заедно с лема 1 показва, че $e(L_2(q))\le 2$.

Нека $G=SL_3(q)$, $\overline{G}=PSL_3(q)$, $q=p^l$. За редовете на групите G и $\overline{G}$ имаме $|G|=q^3(q-1)^2(q+1)(q^2+q+1)$, $|\overline{G}|=d^{-1}|G|$, $d=(3,q-1)$. Групата G съдържа циклични подгрупи от редове $q+1$ и q^2+q+1 и директно произведение на две циклични от редове $q-1$. Ясно е, че ако $d=1$, то единствената силова подгрупа на $\overline{G}$ от нечетен ред и с ранг поне 3 може да бъде само p-групата. Ако $d=3$, то е в сила същото заключение, но трябва да се разгледа отделно силовата 3-подгрупа, тъй като $(q-1,q^2+q+1)=3$. Тя се поражда от диагонални матрици и пермутационна матрица, неабелова е и, както показва непосредственото пресмятане, има ранг 2. Нека $\overline{G}$ съдържа подгрупите A и E от лема 1. Ако $x\in A_0$, $x\ne e$, жордановата му форма трябва да бъде непременно

$$\begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix},$$

защото централизаторът на

$$x'=\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}$$

в $\overline{G}$ е p-група, т. е. x' не централизира инволюции. Като използваме жордановата форма непосредствено пресмятаме, че

$$C=C_{\overline{G}}(x)\cong\left\{\begin{pmatrix}a_{11}&a_{12}&a_{13}\\0&a_{11}&0\\0&a_{23}&a_{33}\end{pmatrix}\mid a_{11}^2a_{33}=1\right\}/Z(G).$$

Следователно $|C_{\overline{G}}(x)|=d^{-1}q^3(q-1)$ и силовата 2-подгрупа на C е циклична, т. е. в C няма четворни групи, което означава, че A централизира инволюция и $|A|=p^3$. От друга страна всички инволюции в $\overline{G}$ са спрегнати с инволюцията $t=diag(-1,-1,1)$ и непосредствено се вижда, че $|C_{\overline{G}}(t)|=|GL_2(q)|d^{-1}=q(q-1)^2(q+1)d^{-1}$ и при $q=p$ или $q=p^2$ подгрупата A не може да се съдържа в $C_{\overline{G}}(t)$, което завършва доказателството на лемата.

За да изследваме групата $L_4(q)$ е необходима значително повече информация за структурата на централизаторите на нейните елементи. За удобство пресмятанията ще

правим в $SL_4(q)$, като непрекъсното ще имаме предвид, че $L_4(q)$ е групата $SL_4(q)$, факторизирана по центъра. Така например, ако $x \in L_4(q)$ и искаме да пресметнем централизатора му, достатъчно е в $SL_4(q)$ да намерим решенията на уравненията

$$y^{-1}xy = \lambda x, \quad y \in SL_4(q), \quad \lambda \in GF(q), \quad \lambda^4 = 1.$$

Ако $(|x|,4)=1$, задължително е $\lambda = 1$.

Лема 5. *Нека* $q = p^l$ *е нечетно,* $G = SL_4(q)$, $\overline{G} = PSL_4(q)$. *Ако*

$$x_1 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$x_3 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad x_4 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

нека $C_i = C_G(x_i)$, $\overline{C}_i = C_{\overline{G}}(x_i)$, $i = 1,2,3,4$. *Тогава*

$$C_1 = \left\{ \begin{pmatrix} \xi_1 & \xi_2 & \xi_3 & \xi_4 \\ 0 & \xi_1 & \xi_2 & \xi_3 \\ 0 & 0 & \xi_1 & \xi_2 \\ 0 & 0 & 0 & \xi_1 \end{pmatrix} | \, \xi_1^{\,4} = 1, \ \xi_2, \xi_3, \xi_4 \in GF(q) \right\}, \quad |C_1| = q^3 d,$$

$|\overline{C}_1| = q^3$, *където тук и по-нататък* $d = (4, q-1)$;

$$C_2 = \left\{ \begin{pmatrix} \xi_1 & \xi_2 & \xi_3 & \alpha \\ 0 & \xi_1 & \xi_2 & 0 \\ 0 & 0 & \xi_1 & 0 \\ 0 & 0 & \beta & \gamma \end{pmatrix} | \, \xi_1^{\,3}\gamma = 1, \ \xi_2, \xi_3, \alpha, \beta \in GF(q) \right\}, \ |C_2| = (q-1)q^4,$$

$|\overline{C}_2| = |C_2|\, d^{-1}$;

$$C_3 = \left\{ \begin{pmatrix} \xi_1 & \xi_2 & \beta_1 & \beta_2 \\ 0 & \xi_1 & 0 & \beta_1 \\ \gamma_1 & \gamma_2 & \alpha_1 & \alpha_2 \\ 0 & \gamma_1 & 0 & \alpha_1 \end{pmatrix} | \, (\xi_1\alpha_1 - \gamma_1\beta_1)^2 = 1, \ \xi_1, \ldots, \gamma_2 \in GF(q) \right\},$$

$$|C_3| = 2\,|SL_2(q)| \cdot q^4 = 2q^5(q^2-1), \quad |\overline{C}_3| = |C_3|\, d^{-1};$$

$$C_4 = \left\{ \begin{pmatrix} \alpha_1 & \alpha_2 & \beta & \gamma \\ 0 & \alpha_1 & 0 & 0 \\ 0 & \delta & a_{11} & a_{12} \\ 0 & \nu & a_{21} & a_{22} \end{pmatrix} | \, \alpha_1^2 \det(a_{ij}) = 1 \right\},$$

$$|C_4| = q^6(q^2-1)(q-1), \quad |\overline{C}_4| = |C_4|\, d^{-1}.$$

Доказателство. Как се пресмята централизаторът на жорданова матрица е добре известно. При $i = 4$ имаме предвид, че $\det(a_{ij}) \neq 0$ трябва да бъде квадрат в $GF(q)$, т.е. за

матрицата (a_{ij}) има $\frac{1}{2}|GL_2(q)|$ възможности. На всяка такава възможност съответстват по две стойности на α_1 (α_1 и $-\alpha_1$).

Лема 6. *Запазваме означенията от лема* 5. *Нека* $q=p$ *е ферматово просто число,* $d=(p-1,4)$ *и нека* S (*съотв.* $\overline{S}$) *означава силова 2-подгрупа на* C_i (*съотв. на* $\overline{C_i}$). *Тогава:*

а) *при* $i=1$ *групата* S *е циклична от ред* d, *а* $|\overline{S}|=1$;

б) *при* $i=2$ *групата* S *е циклична от ред* $p-1$, *а* $\overline{S}$ *е циклична от ред* $(p-1)d^{-1}$;

в) *при* $i=3$: *ако* $p=3$, *то* S *е квазидиедрална група от ред* 16, *а* $\overline{S}$ *е диедрална група от ред* 8; *ако* $p>3$, *то* $\overline{S}$ *е директно произведение на диедрална група от ред* $(p-1)/2$ *и инволюция*;

г) *при* $i=4$, *ако* $p=3$, *то* S *е директно произведение на инволюция и групата на кватернионите от ред* 8, *а* $\overline{S}$ *е групата на кватернионите от ред* 8. *Ако* $p>3$, *то* S *е директно произведение на кватернионна група от ред* $2(p-1)$ *и циклична група от ред* $p-1$, *а* $\overline{S}$ *е директно произведение на кватернионна група от ред* $2(p-1)$ *и циклична група от ред* $(p-1)/4$.

Доказателство. Твърденията а) и б) следват веднага от явното описание на централизаторите в лема 5.

Нека $i=3$, $p=3$ и E означава единичната матрица от ред 2. Нека

$$a=\begin{pmatrix} E & -E \\ E & E \end{pmatrix}, \qquad b=\begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix}.$$

Очевидно $a^8=1$, $a^4=-1$, $b^2=1$, $bab=a^3$ и $a,b\in C_3$ (вж. лема 5). Групата $\overline{S}$ се поражда от образите на a,b при факторизирането.

Нека $i=3$, $p>3$. При $\alpha\in GF(p)$, $\alpha^{p-1}=1$, $\alpha^{\frac{p-1}{2}}=-1$, нека

$$a=diag(\alpha,\alpha,\alpha^{-1},\alpha^{-1}), \quad b=\begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix}, \quad t=\begin{pmatrix} -E & 0 \\ 0 & E \end{pmatrix}.$$

Очевидно $a^{p-1}=b^2=t^2=1$, $b^{-1}ab=a^{-1}$, $b^{-1}tb=-t$. След факторизиране по $Z(G)$ получаваме твърдението.

Нека $i=4$, $p>3$. Тогава матриците

$$a=\begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha^{-1} & 0 \\ 0 & 0 & 0 & \alpha^{-1} \end{pmatrix}, \quad b=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha^{-1} \end{pmatrix}, \quad c=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix},$$

$\alpha^{p-1}=1$, $\alpha^{\frac{p-1}{2}}=-1$, пораждат подгрупата S.

Лема 7. *Ако* p *е просто число на Ферма, групата* $L_4(p)$ *е квазитънка.*

Доказателство. $|L_4(p)|=d^{-1}p^6(p-1)^3(p+1)^2(p^2+p+1)(p^2+1)$, $d=(4,p-1)$. Тъй като $SL_4(p)$ съдържа циклични подгрупи от редове p^2+1 и p^2+p+1, а също и директно произведение на две циклични от редове $p+1$, ясно е, че единствената силова подгрупа от нечетен ред в $L_4(p)$, която би могла да има ранг поне 3, е силовата p-подгрупа. Следователно достатъчно е да докажем, че елементарна абелова подгрупа от ред p^3 не може да нормализира елементарна 2-група.

Нека A и E са подгрупи на $L_4(p)$, които удовлетворяват условието на лема 1. Ще използваме и означенията от лема 5. С точност до спрегнатост $x_i \in A_0$ за някое $i=1,2,3,4$. Следователно A_0, A и E са подгрупи на $\overline{C}_i$. Тъй като $|\overline{C}_1|$ е нечетно число, то $i>2$.

Нека $i=2$. Тогава по лема 6 силовата 2-подгрупа на $\overline{C}_2$ е циклична и съдържа единствена инволюция, следователно A централизира инволюция от $\overline{C}_2$. Тъй като всички инволюции t от $\overline{C}_2$ са спрегнати в $\overline{C}_2$, то с помощта на лема 5 непосредствено пресмятаме, че $|C_{\overline{C}_2}(t)|=p^2(p-1)d^{-1}$. Включването $A\subset C_{\overline{C}_2}(t)$ е невъзможно, защото $|A|=p^3$.

Нека $i>2$ и $p>5$. Сега лема 6 показва, че $\overline{C}_i$ съдържа елементарна 2-подгрупа от ред най-много 8, т. е. $|E|\le 8$. Тъй като $p>5$ е ферматово просто число, то $p\ge 17$ и групата E няма автоморфизъм от ред p. Следователно подгрупата A централизира E. В разглежданата ситуация групата $L_4(p)$ съдържа два класа инволюции с представители

$$t_1=diag(-1,-1,1,1),\quad t_2=diag(\lambda,\lambda,\lambda,\lambda^{-3}),\quad \lambda^8=1,\quad \lambda^4=-1.$$

Централизаторите им в $L_4(p)$ се пресмятат непосредствено. Редовете им са съответно $|GL_2(p)|^2/4(p-1)$ и $|GL_3(p)|/8$. В първия случай силовата p-подгрупа в централизатора е изоморфна на силовата p-подгрупа в $SL_2(p)\times SL_2(p)$, а във втория – на силовата p-подгрупа в $GL_3(p)$. И в двата случая тя има ранг 2, т. е. A не може да се съдържа в централизатор на инволюция.

Нека $i=4,\ p=5$. По лема 6 $|E|\le 4$, следователно подгрупата E няма автоморфизъм от ред 5, т.е. A централизира E. Сега групата $L_4(5)$ има само един клас инволюции с представител $t=diag(-1,-1,1,1,)$ и за реда на централизатора му имаме

$$|C(t)|=\frac{1}{2}p^2(p+1)^2(p-1)^3,$$

следователно p-рангът на $C(t)$ не надминава 2.

Нека $i=3,\ p=5$. Пак по лема 6 $|E|\le 8$. Във всеки от случаите, когато $|E|=2,4$ или 8, подгрупата E няма автоморфизъм от ред 5, следователно $A\subset C(E)$, а редът на централизатора на инволюция не се дели на 5^3.

Нека $i=3,\ p=3$. Сега силовата 2-подгрупа $\overline{S}_2$ в $\overline{C}_3$ е диедрална група от ред 8, следователно редът на E е 2 или 4. Ако с E_2 означим единичната матрица от ред 2, то нека

$$a=\begin{pmatrix} E_2 & -E_2 \\ E_2 & E_2 \end{pmatrix},\quad b=\begin{pmatrix} 0 & E_2 \\ E_2 & 0 \end{pmatrix}.$$

Образите $\overline{a},\overline{b}$ на a,b в групата $L_4(3)$ пораждат силовата 2-подгрупа $\overline{S}_2$, т. е. $\overline{S}_2=<\overline{a},\overline{b}>$. В S_2 има 3 класа инволюции с представители $\overline{a}^2$, $\overline{b}$, $\overline{a}\overline{b}$. Необходимо е да пресметнем техните централизатори в $\overline{C}_3$. За целта е достатъчно при фиксирано t да решим уравненията

$$xt=\pm tx,\quad x\in C_3,\quad t=a^2,\ b,\ ab$$

и да факторизираме множеството на решенията. След непосредствени пресмятания получаваме, че

$$|C_{\overline{C}_3}(\overline{a}^2)|=8\cdot 3^2,\quad |C_{\overline{C}_3}(\overline{b})|=4\cdot 3^2,\quad |C_{\overline{C}_3}(\overline{a}\overline{b})|=8\cdot 3^2.$$

Това показва, че $|E|>2$, защото A не може да централизира инволюция. Следователно $|E|=4$ и всеки елемент $x\in A\setminus A_0$ действа точно върху E. Следователно инволюциите в E са спрегнати в $\overline{C}_3$, откъдето получаваме, че $E\ne<\overline{a}^2,\overline{b}>$, защото централизаторите на $\overline{a}^2$ и $\overline{b}$ в $\overline{C}_3$ имат различни редове. Тъй като всяка диедрална група от ред 8 съдържа точно две

четворни групи, то $E =< \bar{a}^2, \bar{a}\bar{b} >$. Непосредствените пресмятания показват, че $|C_{\bar{C}_3}(E)| = 4 \cdot 3$, докато по предположение $A_0 \subset C_{\bar{C}_3}(E)$ и $|A_0| = 3^2$.

Нека $i = 4,\ p = 3$. Сега силовата 2-подгрупа в $\bar{C}_4$ е кватернионна от ред 8 и съдържа единствена инволюция, следователно A централизира например инволюцията

$$t = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Нейният централизатор в $\bar{C}_4$ е факторгрупата на групата

$$\left\{ \begin{pmatrix} \alpha_1 & \alpha_2 & 0 & 0 \\ 0 & \alpha_1 & 0 & 0 \\ 0 & 0 & a_{11} & a_{12} \\ 0 & 0 & a_{21} & a_{22} \end{pmatrix} | \ \alpha_1^2 \det(a_{ij}) = 1 \right\}$$

по центъра на $SL_4(3)$, редът му очевидно е $2^3 \cdot 3^2$, следователно подгрупата A не може да се съдържа в $C_{\bar{C}_4}(t)$, което завършва доказателството на лемата.

Леми 2, 3, 4, 7 вече дават пълен списък на всички прости квазитънки групи в серията $L_n(q)$ при $n \geq 2$ и произволно q (при $n = 2$, то $q \geq 4$).

ЛИТЕРАТУРА

1. **Aschbaher**, M., Thin finite simple groups, *J. Alg*. 54 (1978), p. 50-152.
2. **Carter**, R., Simple groups of Lie type, London, 1972.
3. **Gorenstein**, D., Finite groups, Harper and Row, New York, 1968.
4. **Gorenstein**, D., Finite simple groups (An introduction to Their Classification, Plenum Press, 1982).
5. **Solomon**, R., A brief history of the classification of the finite simple groups, *Bull AMS*, 38 (2001), 315-352.

ПРОСТИ КВАЗИТЪНКИ ГРУПИ ОТ НОРМАЛЕН И ОТ КРЪСТОСАН ЛИЕВ ТИП

НИКОЛА Т. ПЕТРОВ

SIMPLE QUASI THIN GROUPS OF NORMAL AND TWISTED LIE TYPE

NICOLA T. PETROV

The paper continues the examination, started by the author in [8], what of the known finite simple groups are quasi thin. The main result is stated in [8].

KEY WORDS: finite simples group, quasi thin group

Тази статия е продължение на доказателството на основната теорема, формулирана в статията [8] на автора с подобно заглавие. Запазваме всички означения от [8]. Номерацията на лемите продължава номерацията от [8].

Преминаваме към серията групи ${}^2A_n(q) \cong PSU_{n+1}(q) = U_n(q)$. Известно е, че над крайно поле всеки две неособени ермитови форми са изометрични (вж. Carter [2]) , следователно съответните им унитарни групи са изоморфни. Ако матрицата на използваната неособена ермитова форма е J, то за избягване на недоразуменията съответната специална унитарна или проективна унитарна група понякога ще бележим съответно с $SU_n(q, J)$ и $U_n(q, J)$.

Лема 8. *Групата $U_3(q)$ е квазитънка за всяко четно q.*

Доказателство. Ако $q = 2,$ то $U_3(2)$ има ред 72, очевидно е квазитънка, а освен това е и разрешима. Означаваме

$$J_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad J_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad J_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

По определение

$$SU_3(q, J_i) = \{ X \in SL_3(q^2) \mid {}^t\overline{X} J_i X = J_i \}.$$

Известно е, че

$$|U_3(q)| = d^{-1} q^3 (q-1)(q+1)^2 (q^2 - q + 1), \qquad d = (3, q+1).$$

Групата $SU_3(q, J_2)$ съдържа подгрупата

$$L_1 = \{ diag(\lambda_1, \lambda_2, (\lambda_1 \lambda_2)^{-1} \mid \lambda_i \in GF(q^2),\ \lambda_i^{q+1} = 1 \},$$

която очевидно е изоморфна на директното произведение на две циклични групи от редове $q+1$. Техниката с използване на известната теорема на Ленг за ендоморфизмите на свързаните линейни алгебрични групи ни дава, че $SU_3(q)$ съдържа циклична подгрупа от ред $q^2 - q + 1$. От друга страна $SU_3(q, J_1)$ съдържа подгрупата

$$L_2 = \{ diag(\lambda, \lambda^{q-1}, \lambda^{-q}) \mid \lambda \in GF(q^2),\ \lambda \neq 0 \},$$

която е циклична от ред $q^2 - 1$. Ще забележим, че при четно q имаме : $(q-1, q+1) = 1,$ $(q-1, q^2 - q + 1) = 1$ и $(q+1, q^2 - q + 1) = (3, q+1).$ Следователно всички силови r-подгрупи при $r > 3$ имат ранг най-много 2. Ако $(3, q+1) = 1,$ то 3 дели $q-1$, силовата 3-подгрупа се съдържа в L_2 и следователно е циклична. Ако 3 дели $q+1$, то силовата 3-подгрупа на $SU_3(q, J_2)$ се поражда от силовата 3-подгрупа на L_1 и матрицата

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Тя е неабелова и има ранг 2. От всичко казано следва, че всички силови подгрупи на $U_3(q)$ от нечетен ред имат ранг най-много 2, което доказва твърдението.

Забележка. Ще скицираме и друго доказателство, което не е толкова елементарно, но е по-кратко. Групите $SU_3(q)$ и $U_3(q)$ имат лиев ранг 1, а групата на Вайл W е от ред 2. Като се използва разлагането на Брюа в двойни съседни класове по W лесно се убеждаваме, че силовата 2-подгрупа S_2 на $U_3(q)$ е с тривиално пресичане, т. е. $S_2 \cap S_2^g = \{e\}$ за всеки елемент $g \in U_3(q)$, който е вън от нормализатора на S_2. Следователно всяка 2-локална подгрупа с точност до спрегнатост се съдържа в нормализатора на S_2. От своя страна този нормализатор е разширение на S_2 с помощта на циклична група от ред $d^{-1}(q^2-1),\ d=(3,q+1)$.

Лема 9. *Ако l и n са естествени числа и $5^l = 2^n+1$, то $l=1,\ n=2$.*

Доказателство. Нека $l>1$. Тогава $2^n = 5^l - 1 = 4(5^{l-1}+\ldots+1)$ и тъй като числото в скобите трябва да бъде степен на 2, то l е четно, $l=2k$, и $2^n = (5^k+1)(5^k-1)$, следователно $5^k+1=2^s,\ 5^k-1=2^t$ и $2^s-2^t=2$. Последното равенство е възможно само при $s=2,\ t=1$, откъдето $5^k = 2^t+1=3$, противоречие.

Лема 10. *Ако q е четно и $n>3$, то $e(U_n(q))\geq 3$ с евентуално изключение на групите $U_4(q)$ и $U_5(4)$.*

Доказателство. Нека $n\geq 6$. Ако считаме, че матрицата на съответната ермитова форма е единичната, то групата

$$L = \{diag(\lambda_1,\lambda_2,\lambda_3,\lambda_4,1,\ldots,1) \mid \lambda_i^{q+1}=1,\ \prod\lambda_i = 1\}$$

е от нечетен ред, изоморфна е с подгрупа на $U_n(q)$ и има ранг 3. Тя очевидно централизира инволюция от $U_n(q)$.

Нека $n=5,\ q=2^l,\ l>2$. Съгласно лема 9 числото $q+1$ не е степен на 5 и съществува нечетно просто $p\neq 5$, което дели $q+1$. Нека $\lambda \in GF(q^2),\ \lambda^p=1,\ \lambda\neq 1$. Тогава матриците $a=diag(\lambda,\lambda^{-1},1,1,1),\ b=diag(1,\lambda,\lambda^{-1},1,1),\ c=diag(\lambda^{-2},1,1,\lambda,\lambda)$ са в $SU_5(q)$ и пораждат елементарна група A от ред p^3, която централизира инволюцията

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Тъй като центърът на $SU_5(q)$ има ред 1 или 5, той пресича тривиално подгрупата A и при факторизирането A запазва ранга си. Следователно при $q>4$ имаме $e(U_5(q))\geq 3$. Същото разсъждение може да се повтори и при $q=2$ (тогава $p=3$), следователно $e(U_5(2))\geq 3$.

Лема 11. *Групите $U_4(2^l)$ и $U_5(4)$ са квазитънки за всяко естествено число l.*

Доказателство. $|SU_4(q)| = q^6(q+1)^3(q-1)^2(q^2+1)(q^2-q+1)$. При четно q имаме $(q+1,q-1)=(q+1,q^2+1)=(q-1,q^2-q+1)=1$ и $(q+1,q^2-q+1)=(q+1,3),\ (q-1,q^2+1)=1$. Групата $SU_4(q)$ съдържа циклични подгрупи от редове q^2+1 и q^2-q+1, директно произведение на 3 циклични от ред $q+1$ и директно произведение на две циклични от ред $q-1$. Следователно,

ако p е просто число, по-голямо от 3, то силовата p-подгрупа S_p има ранг най-много 3. Ранг 3 тя може да има само при p, делящо $q+1$. Ако съответната ермитова форма е зададена с единичната матрица, то подгрупата H от ред $(q+1)^3$ се реализира като диагонални матрици:

$$H = \{diag(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \mid \lambda_i^{q+1} = 1,\ \prod \lambda_i = 1\}.$$

Нейният нормализатор в $SU_4(q)$ съдържа всички пермутационни матрици. Нека подгрупите A, A_0, E удовлетворяват условията на лема 1 и $A \subset H$. С точност до спрягане с елементи от $N(H)$ можем да считаме, че $A_0 = < a, b >$, където $a = diag(\mu_1, \mu_2, \mu_3, \mu_4)$, $\mu_i^p = 1$, $\mu_1 \neq 1$, $\prod \mu_i = 1$, $b = diag(1, \nu_2, \nu_3, \nu_4)$, $\nu_i^p = 1$, $\nu_2 \neq 1$, $\prod \nu_i = 1$. Тъй като A има ранг 3, то в $A \setminus A_0$ непременно има матрица от вида $c = diag(1, 1, \tau_3, \tau_3^{-1})$, $\tau_3^p = 1$, $\tau_3 \neq 1$. Вече не е трудно да се съобрази, че всички инволюции, които централизират A, имат вида

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a_{11} & a_{12} \\ 0 & 0 & a_{21} & a_{22} \end{pmatrix}.$$

Но тъй като c нормализира подгрупата E, получаваме, че в $SL_2(q)$ елемент от ред p, където p дели $q+1$, нормализира нетривиална 2-група, което при четно q е невъзможно.

Нека $p = 3$ и 3 дели $q+1$. Нека "построим" силовата 3-подгрупа на $U_4(q)$. Аритметични съображения, свързани с $|U_4(q)|$, показват, че силовата 3-подгрупа на H, дефинирана по-горе, и матрицата

$$d = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

пораждат силовата 3-подгрупа на $U_4(q)$. От друга страна $C_H(d)$ се състои само от матриците $diag(\lambda, \lambda, \lambda, \lambda^{-3})$, $\lambda \neq 0$. Следователно всяка елементарна подгрупа A от ред 3^3 на силовата 3-подгрупа S_3 се съдържа в H и, повтаряйки горното разсъждение с $p = 3$, отново се убеждаваме, че A не се съдържа в 2-локална подгрупа.

Нека 3 дели $q-1$, $q-1 = 3^m \cdot s$, където s не се дели на 3. Тогава $|S_3| = 3^{2m+1}$. Тъй като $U_2(q) \cong SL_2(q)$ при четно q, то $U_4(q)$ съдържа подгрупа, изоморфна на $SL_2(q) \times SL_2(q)$. Такава е например подгрупата

$$\left\{ \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} \mid B, C \in U_2(q) \right\}.$$

Да построим в $U_4(q)$ директно произведение на две циклични групи от ред 3^m (силовите 3-подгрупи на $SL_2(q)$) и да пресметнем централизатора му. Той е изоморфен на $C_{SL_2(q)}(H_0) \times C_{SL_2(q)}(H_0)$, където H_0 е циклична подгрупа на $SL_2(q)$ от ред 3^m. Следователно абеловата подгрупа $H_0 \times H_0$ от ред 3^{2m} съвпада с централизатора си в $U_4(q)$. Впрочем, ако в $H_0 \times H_0$ вземем нециклична подгрупа от ред 9, то непосредствено се проверява, че нейният централизатор също е $H_0 \times H_0$. Следователно абеловият ранг на силовата 3-подгрупа S_3 е 2 и получихме окончателно, че групата $U_4(q)$ е квазитънка.

Преминаваме към групата $U_5(4)$. Тъй като

$$|SU_5(4)| = 5|U_5(4)| = 2^{20}\cdot 3^2\cdot 5^5\cdot 13\cdot 17\cdot 41,$$

достатъчно е да се убедим, че елементарна група A от ред 5^3 не може да нормализира 2-група. В $SU_5(4)$ силовата 5-подгрупа S_5 е разширение на елементарна група от ред 5^4 с помощта на циклична група от ред 5. В "явен вид" тя се поражда от подгрупата H на диагоналните матрици и матрицата

$$a = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

(матрицата на ермитовата форма е единичната матрица). Централизаторът на a в H съвпада с центъра на $SU_5(4)$. Непосредствено се проверява, че ако преминем към $U_5(4)$, то централизаторът на $\bar{a}$ в $\bar{H}$ (образите на a и H при каноничния хомоморфизъм в $U_5(4)$) отново ще има ред 5. Следователно, ако A е групата от лема 1, можем да предполагаме, че е налице включването $A \subset \bar{H}$. Каноничните образи на всички пермутационни матрици от $SL_5(4)$ са в $U_5(4)$ и нормализират $\bar{H}$. С точност до спрягане с тях можем да считаме, че $A_0 = < b, c >$, $b = diag(\lambda_1, \ldots, \lambda_5)$, $\lambda_i^5 = 1$, $\lambda_1 \neq 1$; $c = diag(1, \mu_2, \ldots, \mu_4)$, $\mu_i^5 = 1$, $\mu_2 \neq 1$. Ако x е инволюция от $C(A_0)$, то тя има вида

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & a_{11} & a_{12} & a_{13} \\ 0 & 0 & a_{21} & a_{22} & a_{23} \\ 0 & 0 & a_{31} & a_{32} & a_{33} \end{pmatrix},$$

където матрицата $(a_{ij}) \in SU_3(4)$ и е инволюция в нея. От друга страна, в $A \setminus A_0$ се съдържа матрица $d = diag(1, 1, \lambda, \mu, \nu)$, $\lambda^5 = 1$, $\lambda\mu\nu = 1$. Структурата на нормализатора на силовата 2-подгрупа в $SU_3(4) = U_3(4)$ е добре известна (Gorenstein [3], стр. 466). Той има ред $4^3\cdot 5\cdot 3$, центърът на силовата 2-подгрупа S_2 е от ред 4 и S_2 е с тривиално пресичане в $U_3(4)$. Тъй като d нормализира 2-групата E от лема 1, получаваме, че в $U_3(4)$ някакъв елемент от ред 5 нормализира елементарна 2-група от ред $|E|$, заради тривиалното пресичане нормализира силова 2-подгрупа, следователно нормализира центъра ѝ, той пък е от ред 4, следователно централизира инволюция. За елемента d този извод означава, че елементите λ, μ, ν не са два по два различни. Тъй като b и c централизират инволюцията $x \neq 1$, то $\lambda_3, \lambda_4, \lambda_5$ (съотв. μ_1, μ_2, μ_3) не могат да бъдат два по два различни, защото в противен случай матрицата x ще бъде диагонална. Тъй като всъщност смятаме в $SU_5(4)$, но по модул центъра ѝ, нека забележим, че случаят $\lambda_3 = \lambda_4 = \lambda_5$, $\mu_3 = \mu_4 = \mu_5$ е невъзможен – тогава при факторизирането ще получим група A с ранг по-малък от 3. Нека, например, $\lambda_3 = \lambda_4 \neq \lambda_5$. Тогава задължително $\mu_3 = \mu_4$ и x ще има вида

$$x=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&a_{11}&a_{12}&0\\0&0&a_{21}&a_{22}&0\\0&0&0&0&1\end{pmatrix}.$$

Инволюциите от този вид са в $U_2(4)\cong SL_2(4)$ и 2-групата, която могат да породят е или от ред 2, или четворна. Във всеки от двата случая c е от ред 5 и трябва да централизира x. Ако $\lambda\neq\mu=\nu,$ матрицата x трябва да бъде единичната, следователно $\lambda=\mu$. Но тъй като $\lambda_3=\lambda_4,$ $\mu_3=\mu_4$ и $\lambda=\mu$, отново ще получим матрица, която поражда центъра на $SU_5(4)$ и групата A ще има ранг 2. Случаите $\lambda_3=\lambda_5\neq\lambda_4$ и $\mu_3=\mu_5\neq\mu_4$ чрез спрягане с пермутационна матрица се свеждат до разгледания, което завършва доказателството.

Лема 12. *Нека* $c\in GF(q)$. *Тогава полето* $GF(q^2)$ *е поле на разлагане за полинома* $f(x)=x^q+x+c.$

Доказателство. Следва от очевидното равенство

$$x^{q^2}-x=(f(x))^q-f(x)=f(x)[f(x)^{q-1}-1].$$

Лема 13. *Ако* p *е нечетно просто число, то единственото естествено число* l, *за което* $p^l+1=2^k$, *е числото* $l=1$, *т. е.* p *е мерсеново.*

Доказателство. Тъй като p е най-малко 3, то k е най-малко 2. Ако в горното равенство преминем към сравнение по модул 4 и забележим, че $p\equiv\pm1$ (4), убеждаваме се, че l трябва да бъде нечетно. Нека $l=2l_1+1$ и $l_1\geq1.$ Тогава $(p+1)(p^{2l_1}-p^{2l_1-1}+\ldots+1)=2^k$ и вторият множител от лявата страна е нечетно число, което при $l_1\geq1$ е по-голямо от 1. Следователно $l=1$ и p е мерсеново просто число.

Лема 14. *При нечетно* q *единствените квазитънки групи в серията* $U_n(q)$, $n\geq3$, *са* $U_3(p)$ *и* $U_3(p^2)$ *за всяко нечетно просто число* p *и* $U_4(p)$ *за всяко просто число на Мерсен.*

Доказателство. Най-напред ще се занимаем по-подробно с групата $U_3(q)$. Нейният ред е $d^{-1}q^3(q^2-1)(q^3+1),$ $d=(3,q+1),$ $q=p^l$. Ако реализираме $SU_3(q),$ като използваме формата с матрица J_3 от лема 8, то силовата p-подгрупа S_p на $SU_3(q,J_3)$ ще бъде

$$S_p=\left\{\begin{pmatrix}1&t&u\\0&1&\bar t\\0&0&1\end{pmatrix}\mid t,u\in GF(q^2),\ \bar t=t^q,\ u+\bar u=t\bar t\right\}.$$

Ако за краткост горната матрица означаваме с $(t,u),$ то

$$(t,u)(t',u')=(t+t',u+u'+t\bar t'),$$

откъдето се вижда, че две такива матрици комутират тогава и само тогава, когато $t\bar t'=\bar t t'$. Групата S_p има експонента p, тъй като

$$(t,u)^k=(kt,\ ku+\frac{k(k-1)}{2}t\bar t).$$

Нека $0\neq t\in GF(q^2)$ и да разгледаме подгрупата

$$S_p^{(t)}=\{(\lambda t,u)\mid\ \lambda\in GF(q),\ u+\bar u=\lambda\bar\lambda t\bar t\}.$$

Според казаното тя е елементарна абелова, а от лема 12 следва, че редът ѝ е q^2 (на всеки елемент λt съответстват q стойности на u).

Нека сега да реализираме $SU_5(q)$ с помощта на ермитова форма с единична матрица. Матриците

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad A \in SU_3(q), \quad B \in SU_2(q) \cong SL_2(q)$$

образуват подгрупа, изоморфна на $SU_3(q) \times SL_2(q)$. Тъй като $SU_3(q)$ съдържа елементарна абелова подгрупа от ред q^2, а центърът на $SL_2(q)$ е от ред 2, то очевидно, че $e(U_5(q)) \geq 3$. От това разсъждение се вижда, че $U_n(q)$ не е квазитънка група за всяко $n \geq 5$ и за всяко нечетно q.

Групата $SU_4(q)$ съдържа директното произведение на 3 циклични групи от ред $q+1$, в което се съдържа и центърът й от ред $(4, q+1)$. Следователно за да бъде $U_4(q)$ квазитънка е необходимо $q+1$ да бъде степен на 2, което по лема 13 означава, че $q = p$ и p е просто число на Мерсен. При това условие тя наистина е квазитънка. Редът й е

$$\frac{1}{4} p^6 (p-1)^2 (p+1)^3 (p^2+1)(p^2-p+1).$$

Както в лема 11 се убеждаваме, че е достатъчно да изследваме ранга на p-групите в 2-локалните подгрупи. Разсъжденията провеждаме в $SU_4(p)$ по модул центъра й, който е от ред 4. В групата $U_4(p)$ има точно един клас инволюции с представител $t = diag(-1,-1,1,1)$. Нека $C = C_{U_4(p)}(t)$. Централизаторът C се получава след факторизиране по центъра на $SU_4(p)$ на групата, състояща се от клетъчните матрици от вида

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad \begin{pmatrix} 0 & A_3 \\ A_4 & 0 \end{pmatrix}, \quad A_i \in GL_2(p^2), \quad {}^t\overline{A}_i A_i = E,$$

$$\det(A_1 A_2) = 1, \quad \det(A_3 A_4) = 1, \quad i = 1, \ldots, 4.$$

Силовата p-подгрупа на C има ред p^2 и е елементарна абелова. Ако съществуват подгрупите A и A_0 от лема 1, то A_0 трябва да съвпада (с точност до спрягане) със силова p-подгрупа на C. Но $|C_C(A_0)| = 2$, следователно A централизира инволюция и има ред p^3, което, както видяхме, е невъзможно.

Нека отново $q = p^l$ и да се върнем към групата $U_3(q)$. От лема 12 следва, че групата

$$\{(0,u) \mid u + \overline{u} = 0, \ u \in GF(q^2)\}$$

е елементарна от ред q. Тя централизира инволюцията $diag(-1,1,-1)$, следователно при $l \geq 3$ групата $U_3(q)$ не е квазитънка.

Нека $l = 1$ или $l = 2$. Достатъчно е да се убедим, че абелова p-група с ранг 3 не може да нормализира 2-група. При $l = 1$ това е очевидно, защото силовата p-подгрупа има ред p^3, но не е абелова. Нека $l = 2$. Да забележим, че централизаторът на (t,u) в $U_3(q)$ при $t \neq 0$ е p-група. Следователно, ако подгрупите E и A от лема 1 съществуват, с точност до спрягане

$$A_0 = \{(0,u) \mid u \in GF(p^4), \ u + \overline{u} = 0\}.$$

Сега $(3, p^2+1) = 1$, т. е. $U_3(p^2) = SU_3(p^2)$. Непосредственото пресмятане показва, че $C(A_0)$ не съдържа четворни групи, т. е. A централизира инволюции. Това обаче е невъзможно, защото A ще съдържа матрици (t,u) с $t \neq 0$, които не централизират инволюции. Лемата е доказана, с което завършихме изследването на серията унитарни групи.

Преминаваме към серията $B_n(q)$. При четно q ще използваме, че $B_n(q) \cong C_n(q) \cong PSp(2n,q)$.

Оттук нататък свободно и без пояснения ще използваме общоприетата терминология и означения в теорията на групите на Шевалие. Всички означения и терминологията може да се намерят в лекциите на Стайнберг [4].

Лема 15. *При четно* q *единствените квазитънки прости групи в серията* $B_n(q)$, $n \geq 2$, *са групите* $B_2(q)$ *и* $B_3(2)$, *където* $q > 2$ *е произволна степен на числото* 2.

Доказателство. При четно q подгрупата на Картан H в $B_n(q)$ е директно произведение на n циклични групи от редове $q-1$ и нормализира силовата 2-подгрупа. Следователно при $n \geq 3$ и $q > 2$ групата $B_n(q)$ не е квазитънка. Лесно е да се съобрази, че при $n \geq 4$ групата $B_n(q)$ съдържа директно произведение на група от ред 2 и 3 екземпляра на групата $SL_2(q)$. Наистина, нека

$$J = \begin{pmatrix} 0 & 1 & & & & & \\ -1 & 0 & & & & & \\ & & 0 & 1 & & & \\ & & -1 & 0 & & & \\ & & & & \cdots & & \\ & & & & & 0 & 1 \\ & & & & & -1 & 0 \end{pmatrix}.$$

По определение

$$Sp(2n, q) = \{X \in GL_{2n}(q) \mid {}^tXJX = J\}$$

за всяко четно или нечетно q. Известно е, а и лесно се проверява, че е налице равенството $SL_2(q) = Sp(2, q)$. Следователно клетъчната матрица

$$\begin{pmatrix} A_1 & & & \\ & A_2 & 0 & \\ & 0 & \ddots & \\ & & & A_n \end{pmatrix}, \quad A_i \in SL_2(q),$$

е симплектична и горното твърдение става очевидно. Остава да разгледаме само групите $B_2(q)$ и $B_3(2)$. Групата $B_2(q)$ съдържа по-специално циклична подгрупа от ред q^2+1, директно произведение на две циклични групи от редове $q-1$ и директно произведение на две циклични от редове $q+1$. Тъй като

$$|B_2(q)| = q^4(q-1)^2(q+1)^2(q^2+1),$$

ясно е, че всички силови подгрупи от нечетен ред имат ранг най-много две, следователно $B_2(q)$ е квазитънка. (Групата $B_2(2) \cong S_6$ не е проста.)

Редът на групата $B_3(2)$ е $2^9 \cdot 3^4 \cdot 5 \cdot 7$, следователно интерес представляват само 3-подгрупите. Нека $a = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, $e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Силовата 3-подгрупа S_3 се поражда от клетъчно-диагоналните матрици от вида $diag(a^i, a^j, a^k)$, $1 \leq i, j, k \leq 3$ и клетъчната матрица

$$Q = \begin{pmatrix} 0 & 0 & e \\ e & 0 & 0 \\ 0 & e & 0 \end{pmatrix}.$$

Ясно е, че групата S_3 е неабелова, съдържа нормална елементарна подгрупа от ред 3^3, а центърът ѝ е от ред 3. Следователно S_3 съдържа единствена елементарна подгрупа от ред 27,

която ще играе ролята на подгрупата A от лема 1, т. е. $A = \{ diag(a^i, a^j, a^k), 1 \le i, j, k \le 3 \}$. Нека A, A_0 и E са подгрупите от лема 1. С точност до спрягане с матрицата Q можем да считаме, че подгрупата A_0 се поражда от клетъчно-диагоналните матрици $diag(a, e, a^l)$, $diag(e, a, a^m)$, където l, m са подходящи фиксирани числа, $1 \le l, m \le 3$. Но тогава $C(A_0)$ се състои от някакви клетъчно-диагонални матрици, в които клетките по диагонала са матрици от $SL_2(2)$. Ако е налице поне едно от неравенствата $l \ne 3, m \ne 3$, то $C(A_0) = A$, следователно подгрупата A_0 не централизира инволюции. Ако $l = m = 3$, то имаме $C(A_0) \cong < a > \times < a > \times SL_2(2) = A_0 \times SL_2(2)$. Сега, ако подгрупата A нормализира 2-групата E, то A ще нормализира и сечението $C(A_0) \cap E$. Силовата 2-подгрупа в $C(A_0)$ е от ред 2, следователно сечението или се състои само от единицата, и тогава A_0 не централизира инволюция от E, или сечението има ред 2, а тогава подгрупата A централизира инволюция от $C(A_0)$. Това е невъзможно, защото в $SL_2(2)$ централизаторът на всяка инволюция е от ред 2. Лемата е доказана.

Лема 16. *При нечетно $q = p^l$ единствените квазитънки групи в серията $B_n(q), n > 1$, са групите $B_2(p)$ за всяко нечетно просто число p.*

Доказателство. Ако α, β са корени на съответната алгебра на Ли, то

$$h_\alpha(t) x_\beta(u) h_\alpha(t)^{-1} = x_\beta(t^{<\beta,\alpha>} u),$$

където $< \beta, \alpha > = 2(\alpha, \beta)/(\alpha, \alpha)$. Като използваме означенията от Бурбаки [7], стр. 304, означаваме

$$A_1 = \left\langle x_{\alpha_{n-1}}(t), x_{\alpha_{n-1}+2\alpha_n}(u) \mid t, u \in GF(q) \right\rangle.$$

Сумата на корените α_{n-1} и $\alpha_{n-1} + 2\alpha_n$ не е корен, следователно подгрупата A_1 е елементарна абелова от ред q^2. Елементът $h_{\alpha_{n-1}}(-1)$ не е единица, а инволюция, която централизира A_1. Наистина, $< \alpha_n, \alpha_{n-1} > = -1$, $< \alpha_{n-1} + \alpha_n, \alpha_{n-1} > = 1$, $< \alpha_{n-1} + 2\alpha_n, \alpha_{n-1} > = 0$, $< \alpha_{n-1}, \alpha_{n-1} > = 2$ и формулата по-горе ни убеждава в казаното. Следователно при $l > 1$ групата $B_n(p^l)$, $n \ge 2$, не е квазитънка.

Да предположим, че $n \ge 3$ и нека

$$A_2 = \left\langle x_{-\alpha_{n-1}}(t),\ x_{\alpha_{n-1}+2\alpha_n}(t),\ x_{\alpha_{n-2}+\alpha_{n-1}+\alpha_n}(t) \mid t \in GF(q) \right\rangle.$$

Отново сумата на никои два от използваните корени не е корен, следователно A_2 е елементарна абелова от ред q^3 (рангът е $3l$), която се централизира от инволюцията $h_{\alpha_{n-1}}(-1)$. (За да се убедим в това е достатъчно да погледнем матрицата на Картан на системата корени от тип B_n (вж. Бурбаки [7].) Този факт показва в частност, че ако $n \ge 3$ и $q = p$ е нечетно просто число, то групата $B_n(p)$ не е квазитънка.

Сега ще покажем, че $B_2(p)$ е квазитънка. Тъй като

$$| B_2(p) | = \frac{p^4}{2} (p-1)^2 (p+1)^2 (p^2+1),$$

същите съображения както в лема 15 показват, че е достатъчно да изследваме само p-групите в 2-локалните подгрупи. Тъй като е налице изоморфизмът $B_2(p) \cong PSp(4, p)$, ще използваме информацията от [6]. В $B_2(p)$ има не повече от два класа инволюции. Първият е с представител $t_1 = diag(1, 1, -1, -1)$ и централизаторът му е изоморфен с централно произведение на $SL_2(p)$ с $SL_2(p)$. Вторият се получава от елементите t_2 на $Sp(4, p)$, чиито квадрати са инволюцията от центъра на $Sp(4, p)$. В означенията от [6], ако 4 дели $p - 1$, такъв е класът

$B_8(\frac{p-1}{4})$, а ако 4 дели $p+1$ - класът $B_6(\frac{p+1}{4})$. За инволюция от втория клас централизаторът ѝ е или от ред $\frac{p}{2}(p-1)(p^2-1)$, или от ред $\frac{p}{2}(p+1)(p^2-1)$, следователно не се дели на p^2. И така, ако в $B_2(p)$ съществуват подгрупите A и E от лема 1, то E съдържа инволюции само от първия клас. Можем да считаме, че A_0 съвпада със силова p-подгрупа на централизатора $C_1 = C(t_1)$. Но $C_{C_1}(A_0)$ съдържа единствена инволюция, следователно A я централизира. Но това е невъзможно, защото p^3 не дели реда на централизатор на инволюция.

Лема 17. *При нечетно q и $n \geq 3$ групите от серията $C_n(q)$ не са квазитънки.*

Доказателство. Нека

$$A = \left\langle x_{\alpha_n}(t),\ x_{2\alpha_{n-1}+\alpha_n}(t),\ x_{\alpha_{n-2}+2\alpha_{n-1}+\alpha_n}(t) \mid t \in GF(q) \right\rangle.$$

Тъй като сумата на никои два от използваните корени не е корен, то A е елементарна абелова група с ранг поне 3, а справка с Бурбаки [7], стр. 306, показва, че $h_{\alpha_n}(-1)$ е инволюция, която централизира A.

Забележка. Тъй като при четно q е налице изоморфизмът $B_n(q) \cong C_n(q)$, вече не е нужно да изследваме групите от тип C_n при четно q.

Лема 18. *В серията групи $D_n(q)$, $n \geq 4$, няма квазитънки групи каквото и да бъде q.*

Доказателство. Нека $A_i = \left\langle x_{r_i}(t),\ x_{-r_i}(t) \mid t \in GF(q) \right\rangle$, $i = 1,\ldots,4$, като в означенията на Бурбаки [7], стр. 308, $r_1 = \varepsilon_1 + \varepsilon_2$, $r_2 = \varepsilon_1 - \varepsilon_2$, $r_3 = \varepsilon_3 + \varepsilon_4$, $r_4 = \varepsilon_3 - \varepsilon_4$. Всяка от групите A_i е хомоморфен образ на групата $SL_2(q)$, като ядрото се съдържа в центъра на $SL_2(q)$. Ако q е четно, ядрото е тривиално. Ще покажем, че то е тривиално и при нечетно q. Достатъчно е да се убедим, че за всяко $i = 1,\ldots,4$ имаме $h_{r_i}(-1) \neq 1$. Това се проверява непосредствено: за всеки корен r_i е лесно да се посочи корен r_i', за който $< r_i', r_i >= \pm 1$. Корените r_i са два по два ортогонални, следователно A_i централизира A_j при $i \neq j$. Всичко това означава, че групата $D_n(q)$ съдържа директното произведение на четири екземпляра на групата $SL(2,q)$, което доказва лемата.

Лема 19. *Сериите групи $E_i(q)$, $i = 6,7,8$, не съдържат квазитънки групи.*

Доказателство. Ако q е четно, универсалната група на Шевалие $D_4(q)$ съвпада с присъединената и се влага изоморфно в $E_i(q)$. Тъй като $D_4(q)$ не е квазитънка, такава е и $E_i(q)$. При нечетно q в $E_i(q)$ по очевиден начин се строи директно произведение на три екземпляра на групата $SL_2(q)$, което доказва твърдението и в този случай.

Лема 20. *Групите $F_4(q)$ не са квазитънки.*

Доказателство. Групата на Шевалие $F_4(q)$ съдържа хомоморфен образ на универсалната група на Шевалие $D_4(q)$, а тя, както и хомоморфните ѝ образи, не е квазитънка.

ЛИТЕРАТУРА

1. **Borel**, A., J. Tits, Elements unipotents et sous-groupes paraboliques des groupes reductifs, *Invent. Math.* **12**, (1971), 95- 104.
2. **Carter**, R., Simple groups of Lie type, Wiley-Interscience, London, 1972.
3. **Gorenstein**, D., Finite groups, Harper and Row, New York, 1968.
4. **Steiberg**, R., Lectures on Chevalley groups, Lectures Notes, Yale University, 1967-1968,
5. **Steiberg**, R., Endomorphisms of linear algebraic groups, Mem. Amer. Math. Soc. 1968, 80, 108 pp.
6. **Srinivasan**, B., The characters of the finite symplectic groups Sp(4, q), *Trans. A.M.S* 131, (1968), 488-525.
7. **Бурбаки**, Н., Группы и алгебры Ли, гл. IV-VI, М., Мир, 1972.
8. **Петров**, Н., Прости квазитънки групи от нормален лиев тип, Сборник научни трудове посветен на 10-годишнината от създаването на Факултет по математика и информатика, УИ „Епископ К. Преславски”, 2007.

ПРОСТИ КВАЗИТЪНКИ ГРУПИ ОТ КРЪСТОСАН ЛИЕВ ТИП И СПОРАДИЧНИ КВАЗИТЪНКИ

НИКОЛА Т. ПЕТРОВ

SIMPLE QUASI THIN GROUPS OF TWISTED LIE TYPE AND SPORADIC QUASI THIN GROUPS

NICOLA T. PETROV

The paper completes the examination started by the author in [7] and [8] what of the known finite simple groups are quasi thin. The main result is stated in [7].

KEY WORDS: finite simples group, quasi thin group

Тази статия е продължение на статията [8] и завършва доказателството на основната теорема, формулирана в статията на автора [7]. Запазваме всички означения от [7] и [8]. Номерацията на лемите продължава номерацията от [8].

Лема 21. *Единствените квазитънки групи в серията групи от тип* G_2 *са* $G_2(q)$ *за всяко четно* q *и* $G_2(p)$ *за всяко просто число* p.

Доказателство. Известно е, че

$$|G_2(q)| = q^6(q-1)^2(q+1)^2(q^2+q+1)(q^2-q+1).$$

За всяко q групата $G_2(q)$ съдържа циклични подгрупи от редове q^2+q+1 и q^2-q+1, директно произведение на две циклични от редове $q-1$ и директно произведение на две циклични от редове $q+1$. Тази информация и аритметични съображения, свързани с реда на $G_2(q)$ показват, че при четно q всяка силова p - подгрупа на $G_2(q)$ при $p>3$ има ранг най-много 2. Нека $p=3$ и q е четно. Ако 3 дели $q-1$, то 3 дели и q^2+q+1, но 9 не дели q^2+q+1. Следователно силовата 3-подгрупа на $G_2(q)$ е силовата 3-подгрупа на картановата подгрупа H, разширена с елемент w, $|w|=3$, от подгрупата на Вайл W. От това се вижда, че тя е неабелова, а рангът й е две. До същия извод се стига и ако забележим, че корневите подгрупи, съответстващи на дългите корени в алгебрата на Ли от тип G_2, пораждат $SL_3(q)$ и силова 3-подгрупа на $G_2(q)$ се съдържа в подгрупа $SL_3(q)$. Следователно, ако 3 дели $q-1$ и q е четно, то $G_2(q)$ е квазитънка. Ако 3 дели $q+1$, то 3 дели q^2-1 и според доказаното групата $G_2(q^2)$ е квазитънка. Тя съдържа $G_2(q)$ и следователно последната също е квазитънка.

Забележка. В частност квазитънка е групата $G_2(2)$, но тя не е проста. Нейният комутант е проста група и следователно $G_2(2)'$ е квазитънка група.

Нека $q=p^l$, където p е нечетно просто число. Тъй като групата $G_2(q)$ съдържа централното произведение на $SL_2(q)$ със $SL_2(q)$, то при $l\geq 2$ тя очевидно не е квазитънка.

Ще покажем, че при нечетно просто p групата $G_2(p)$ е квазитънка. За целта ще използваме информация от статиите на Chang [2] и Enomoto [3]. Както по-горе се убеждаваме, че е достатъчно да изследваме само p -групите в 2-локалните подгрупи. Групата $G_2(p)$ съдържа само един клас инволюции (в означенията от [2] и [3] представител на класа е $h(-1,-1,1)$). Ако t е инволюция, то

$$|C(t)| = p^2(p^2-1)^2.$$

Следователно p-групата A от лема 1 не може да централизира инволюция, но $A_0 \subset C(t)$ и в $C(t)$ се съдържат четворни групи. Последното е невъзможно: справка с [2] и [3] показва, че силовите 2-подгрупи в централизатор на p-елемент или са циклични от ред ≤ 2, или са кватернионни, следователно не съдържат четворни групи. Лемата е доказана.

Лема 22. *Групите на* Сузуки $Sz(q)$, $q = 2^{2l+1}$, *са тънки.*

Доказателство. Структурата на групите на Сузуки е добре известна (вж. [5]). В тях всички силови подгрупи от нечетен ред са циклични.

Лема 23. *Групите на* Ри ${}^2G_2(3^{2l+1})$, $l \geq 1$, *не са квазитънки.*

Доказателство. В групите на Ри има само един клас инволюции и ако t е инволюция, то

$$C(t) \cong < t > \times PSL_2(3^{2l+1}).$$

Ясно е, че централизаторът на t съдържа елементарна група от ред 27.

Лема 24. *В серията алтернативни групи* A_n, $n \geq 5$, *квазитънки са само групите* A_5, A_6, A_7, A_8, A_9.

Доказателство. В групата A_{10} по очевиден начин се построява директно произведение на A_4 с елементарна група от ред 9. Така построената група нормализира силовата си 2-подгрупа, следователно A_{10} не е квазитънка. Тъй като при $n \geq 10$ имаме $A_{10} \subset A_n$, то алтернативните групи A_n не са квазитънки за всяко $n \geq 10$.

За групите A_i, $i = 5,6,7,8$ твърдението в лемата е очевидно, защото силовите им p-подгрупи от нечетен ред имат ред най-много p^2, следователно рангът им не надминава 2.

Да разгледаме групата A_9. Тя има точно два класа инволюции с представители $t_1 = (1\,2)(3\,4)$ и $t_2 = (1\,2)(3\,4)(5\,6)(7\,8)$. Централизаторът $C(t_1)$ (в A_9) е директно произведение на алтернативната група A_5 с четворна група, разширено с инволюцията (12)(56). Централизаторът $C(t_2)$ (в A_9) се съдържа в A_8 и редът му не се дели на 9, защото в противен случай ще съдържа произведение на два независими тройни цикъла, което е невъзможно. И в двата случая не съществува нечетно просто число, чийто квадрат да дели реда на централизатора на инволюция, което по лема 1 означава, че групата A_9 е квазитънка.

Лема 25. *Групите* ${}^2D_n(q)$, $n \geq 4$, *не са квазитънки.*

Доказателство. Отново без уговорки ще използваме означенията от Бурбаки [6], стр. 308. Ще забележим, че за да намерим образа на който и да е корен при симетрията на схемата на Динкин, достатъчно е само да разменим ε_n и $-\varepsilon_n$. Нека r е корен, а $\bar{r}$ е образът му при симетрията на схемата на Динкин. Ако $r = \bar{r}$, то $\langle x_r(t) \mid t \in GF(q) \rangle \subset {}^2D_n(q)$. Следователно при $n > 4$ групата $A_1 \times A_2 \times A_3 \times A_4$ от доказателството на лема 18 се съдържа в ${}^2D_n(q)$, а това в случая доказва твърдението. При $n = 4$ ще забележим, че $x_{\pm r_3}(t)x_{\pm r_4}(\bar{t}) \in {}^2D_2(q)$ и групата

$$A_0 = \left\langle x_{r_3}(t)x_{r_4}(\bar{t}),\ x_{-r_3}(t)x_{-r_4}(\bar{t}) \mid t \in GF(q^2) \right\rangle$$

е изоморфна на $L_2(q^2)$, защото сега $h_{r_3}(-1)h_{r_4}(-1) = 1$. Тъй като A_1 и A_2 отново са подгрупи на ${}^2D_4(q)$, получаваме, че ${}^2D_4(q)$ съдържа подгрупа, изоморфна на

$$SL_2(q) \times SL_2(q) \times L_2(q),$$

което завършва доказателството на лемата, защото последната група не е квазитънка.

Лема 26. *При нечетно* q *групата* ${}^3D_4(q)$ *не е квазитънка. При четно* q *тя е квазитънка.*

Доказателство. Нека $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ са система прости корени от тип D_4 и нека α_2 е свързан с всеки от останалите. Ако $t\in GF(q^3)$, нека $t^q=\bar{t}$. При тези означения групата

$$L=\left\langle x_{\alpha_1}(t)x_{\alpha_3}(\bar{t})x_{\alpha_4}(\bar{\bar{t}})\ \ x_{-\alpha_1}(t)x_{-\alpha_3}(\bar{t})x_{-\alpha_4}(\bar{\bar{t}})\,|\,t\in GF(q^3)\right\rangle$$

е изоморфна на $SL_2(q^3)$ и се съдържа в ${}^3D_4(q)$. Тъй като при нечетно q групата $SL_2(q^3)$ не е квазитънка, такава е и ${}^3D_4(q)$.

Нека q е четно, а ρ е симетрията от ред 3 на схемата на Динкин; ако $\rho(\alpha_1)=\alpha_3$, то ρ-орбитите в множеството на положителните корени са:
$\{\alpha_2\}$, $\{\alpha_1+\alpha_2+\alpha_3+\alpha_4\}$, $\{\alpha_1+2\alpha_2+\alpha_3+\alpha_4\}$, $\{\alpha_1+\alpha_2,\alpha_3+\alpha_2,$
$\alpha_4+\alpha_2\}$, $\{\alpha_1+\alpha_3+\alpha_2,\alpha_3+\alpha_4+\alpha_2,\alpha_4+\alpha_1+\alpha_2\}$.

Групата ${}^3D_4(q)$ има лиев ранг 2, а групата й на Вайл е изоморфна на групата на Вайл на алгебрата от тип G_2, т.е. диедрална от ред 12. Силовата 2-подгрупа на ${}^3D_4(q)$ има ред q^{12}, а 2-допълнението в нормализатора й B е от ред $(q-1)(q^3-1)$ (директно произведение на две циклични групи съответно от редове $q-1$ и q^3-1). Съгласно теоремата на Борел и Титс всяка 2-локална подгрупа с точност до спрягане се съдържа в максимална параболична подгрупа. В разглежданата група максималните параболични подгрупи (с точност до спрягане) са точно две:

$$P_1=\langle B,w_2\rangle,\quad P_2=\langle B,w_1w_3w_4\rangle,$$

където $w_i=w_{\alpha_i}$ е отражението относно хиперравнината, перпендикулярна на корена α_i. Тъй като ρ пермутира положителните корени, различни от α_2, то $O_2(P_1)$ е от ред q^{11} и $P_1/O_2(P_1)$ е полудиректно произведение на $SL_2(q)$ със циклична група от ред q^3-1. Следователно P_1 е квазитънка. Със същото разсъждение получаваме, че $P_2/O_2(P_2)$ е полудиректно произведение на $SL_2(q^3)$ със циклична група от ред $q-1$, т. е. параболичната група P_2 също е квазитънка. Както вече отбелязахме, по теоремата на Борел и Титс това означава, че групата ${}^3D_4(q)$ е квазитънка, което завършва доказателството на лемата.

Лема 27. *Групите* ${}^2F_4(2^{2l+1})$ *са квазитънки за всяко* $l\geq 0$. *Групата* ${}^2F_4(2)'$ *е тънка.*

Доказателство. В групата ${}^2F_4(2)$ 2-локалните подгрупи имат редове,делящи $2^{12}\cdot 3$ и $2^{12}\cdot 5$, следователно тя даже е тънка. Тъй като комутантът ${}^2F_4(2)'$ има индекс 2 в ${}^2F_4(2)$, той също е тънка група.

Нека $l\geq 1$. Групата ${}^2F_4(q)$ има лиев ранг 2, следователно има точно две максимални параболични подгрупи P_1 и P_2. С непосредствени пресмятания се вижда, че $P_1/O_2(P_1)$ е изоморфна на полудиректно произведение на $Sz(q)$ с циклична група от ред $q-1$, а $P_2/O_2(P_2)$ е изоморфна на полудиректно произведение на $SL_2(q)$ с циклична група от ред $q-1$. По теоремата на Борел и Титс това означава, че ${}^2F_4(q)$ е квазитънка, тъй като при четно q групите $Sz(q)$ и $SL_2(q)$ са тънки.

Лема 28. *Групите* ${}^2E_6(q)$ *не са квазитънки.*

Доказателство. Достатъчно е да забележим, че групата ${}^2E_6(q)$ съдържа хомоморфен образ на групата $SU_6(q)$, а предишните разглеждания показват, че нито $SU_6(q)$, нито хомоморфните й образи са квазитънки.

Лема 29. *Между* 26*-те спорадични групи квазитънки са само следните*: M_{11}, M_{12}, M_{22}, M_{23}, M_{24}, J_1, J_2, J_3, J_4, HS, He, Ru, M^C.

Доказателство. Значителна част от необходимата информация може да се намери в обзора на Сыскин [9]. Петте групи на Матийо са квазитънки, защото или няма нечетно просто число, чийто куб да дели реда на групата (такива са M_{11}, M_{22}, M_{23}), или ако има такова число, то е 3, силовата 3-подгрупа е от ред 27 и е неабелова. Същото важи и за групите на Янко J_1 и J_2, които имат редове $2^3 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19$ и съответно $2^7 \cdot 3^3 \cdot 5^2 \cdot 7$. За групата J_2 бихме могли да се позовем и на факта, че тя се влага в $G_2(4)$, за която доказахме, че е квазитънка. В групата на Янко J_4 силовите p-подгрупи от нечетен ред или са от ред p, или са от ред 3^3 и 11^3, като в последния случай са неабелови. За групата J_3 разполагаме с класификация на максималните й подгрупи (вж. [4]). Всички те са квазитънки.

Групата на Конуей Co_2 съдържа максимална подгрупа, изоморфна на разширение на $U_6(2)$ с помощта на инволюция. Тъй като $U_6(2)$ не е квазитънка, то и Co_2 не еквазитънка. Групата Co_2 се влага в групите Co_1, F_1 и F_2, следователно те също не са квазитънки.

Групата на Конуей Co_3 съдържа инволюция z, чийто централизатор е неразцепимо разширение на $<z>$ с помощта на $Sp_6(2)$. Както видяхме в предишни леми, групата $Sp_6(2)$ съдържа елементарна абелова подгрупа от ред 27, следователно Co_3 не е квазитънка.

Групата на Фишер F_{22} съдържа инволюция d и $C(d)/<d> \cong U_6(2)$. Както видяхме, $U_6(2)$ не е квазитънка, следователно F_{22} също не е квазитънка. Тъй като F_{22} се влага в групите на Фишер F_{23} и F_{24}', те също не са квазитънки.

Групата на Хигман – Симс HS има ред $2^9 \cdot 3^2 \cdot 5^3 \cdot 7 \cdot 11$. Тъй като силовата й 5-подгрупа не е абелова, тя е квазитънка. Аналогично съображение е приложимо и за групата на Хелд He, която има ред $2^{10} \cdot 3^3 \cdot 5^2 \cdot 7^3 \cdot 17$.

Групата на Сузуки Suz съдържа 2-локална подгрупа, която е разширение на група от ред 2^7 (централно произведение на 3 екземпляра кватерниони от ред 8) с помощта на $PSp_4(3)$. Тъй като $PSp_4(3)$ съдържа елементарна подгрупа от ред 27, то Suz не е квазитънка.

Групата на Маклафлин Mc е квазитънка, защото съдържа точно 12 класа максимални подгрупи и всички те са квазитънки.

Групата на Лайонс Ly не е квазитънка, защото съдържа централна инволюция, чийто централизатор е изоморфен на неразцепимо разширение на инволюцията с алтернативната група A_{11}.

Групата на Рудвалис Ru има ред $2^{14} \cdot 3^3 \cdot 5^3 \cdot 7 \cdot 13 \cdot 29$. Силовата 3-подгрупа и силовата 5-подгрупа са неабелови, а останалите силови подгрупи от нечетен ред са циклични. Следователно Ru е квазитънка.

В групата на О'Нан ON има елемент от ред 3, чийто централизатор C е директно произведение на елементарна група от ред 9 с $L_2(9)$. Тъй като в групата $L_2(9)$ има елемент от ред 3, който нормализира четворна подгрупа, то C не е квазитънка група , а следователно и ON не е квазитънка.

В групата на Томпсън F_3 централизаторът на всяка инволюция е разширение на група от ред 2^9 с помощта на алтернативната група A_9. Тъй като A_9 съдържа елементарна група от ред 27, то F_3 не е квазитънка.

Групата на Харада F_5 не е квазитънка, защото съдържа разширение на елементарна група от ред 2^6 с $PSp_4(3)$.

Разгледахме всичките 26 спорадични групи и лемата е доказана. По този начин завършихме и проверката кои от известните прости крайни групи са квазитънки, а това завършва доказателството на теоремата, формулирана в [7].

ЛИТЕРАТУРА

1. **Borel**, A., J. Tits, Elements unipotents et sous-groupes paraboliques de groupes reductifs: I. *Invent. Math.*, 1971, 12, 95-104.
2. **Chang**, B., The conjugacy classes of Chevalley groups of type (G_2), *J. Alg.*, 9, (1968), 190-211.
3. **Enomoto**, H., The conjugacy classes of Chevalley groups of type (G_2) over finite fields of characteristic 2 or 3, *J. Fac. Sci. Univ. Tokyo,* 16, (1970), 497-512.
4. **Finkelstein**, L., A. Rudvalis, The maximal subgroups of the Janko's simple group of order 50 232 960, *J. Alg*. 30, (1974), p. 133-143.
5. **Suzuki**, M., On a class of doubly transitive groups, I, II, *Ann. Math.*, 1962, 75, 105-145; 1964, 79, 514-589.
6. **Бурбаки**, Н., Группы и алгебры Ли, гл. IV-VI, М., Мир, 1972.
7. **Петров**, Н., Квазитънки групи от нормален лиев тип, Сборник научни трудове посветен на 10-годишнината от създаването на Факултет по математика и информатика, УИ „Епископ К. Преславски", 2007.
8. **Петров**, Н., Квазитънки групи от нормален и кръстосан лиев тип, Сборник научни трудове посветен на 10-годишнината от създаването на Факултет по математика и информатика, УИ „Епископ К. Преславски", 2007.
9. **Сыскин**, С. А., Абстрактные свойства простых спорадических групп, *УМН*, т. 35, вып. 5, (1980), 181-212.

ВЪРХУ НЯКОИ ПЕРТУРБАЦИОННИ ОЦЕНКИ НА МАТРИЧНИТЕ УРАВНЕНИЯ $X + A^* X^{-1} A = Q$ И $X - A^* X^{-1} A = Q$

ВЕЖДИ И. ХАСАНОВ, АЙНУР А. ФЕЙЗОЛОВА

ON SOME PERTURBATION ESTIMATES OF THE MATRIX EQUATIONS $X + A^* X^{-1} A = Q$ AND $X - A^* X^{-1} A = Q$

VEJDI I. HASANOV, AYNUR A. FEYZOLOVA

Some perturbation estimates for positive definite solutions of two nonlinear matrix equations are considered. Using numerical examples we compare different estimates concerning the true error.

KEY WORDS: Matrix equations, positive definite solutions, perturbation estimates

1. Въведение

За дискретна линейна стационарна управляваща система задачата за оптималното решение се свежда до решаване на дискретното уравнение на Рикати, а именно

$$X - S^* XS + S^* XB\left(R + B^* XB\right)^{-1} B^* XS - Q = 0\,. \tag{1}$$

Следвайки Engwerda [1] и Ferrante и Levy [2] решаването на уравнението (1) при специални мръзки между коефициентите се свежда до решаване на уравненията

$$X + A^* X^{-1} A = Q \tag{2}$$

и

$$X - A^* X^{-1} A = Q \tag{3}$$

съответно.

Уравненията (2) и (3) за съществуване на положително определени решения са изследвани от много автори, като едни от първите са Engwerda [1] и Ferrante и Levy [2]. Едно положително определено решение X_L на матрично уравнение наричаме максимално, ако за всяко друго ермитово решение X е изпълнено $X \le X_L$.

В настоящата работа са разгледани някои известни пертурбационни оценки, които са сравнени по отношение на близост до истинската грешка с няколко числени примери.

Използвани са: $\|\cdot\|$-спектрална норма; $\|\cdot\|_F$-норма на Фробениус и $\|\cdot\|_U$-произволна унитарна норма.

2. Пертурбационни оценки

Тук са дадени някои известни пертурбационни оценки за разгледаните уравнения.

Преди да разгледаме двете уравнения по отделно ще представим две оценки за решенията на уравнение $X + A^* F(X) A = Q$ [4], които при смяна на матричната фунция $F(X)$ съответно с X^{-1} и $-X^{-1}$ са приложими за уравненията (2) и (3).

Теорема 1. *(Proposition 4.1 [4]) Нека X и $\widetilde{X}$ са положително определени решения в S_n на матричните уравнения $X + A^* F(X) A = Q$ и $\widetilde{X} + \widetilde{A}^* F\left(\widetilde{X}\right)\widetilde{A} = \widetilde{Q}$ съответно. Ако $M_{S(n)}\|A\|^2 < 1$ и е изпълнено:*

$$\left\|\tilde{A}\right\| \le \|A\| \quad или \quad \left\|\tilde{A}\right\| > \|A\| \quad и \quad \left\|A-\tilde{A}\right\| \le \frac{1-M_{S(n)}\|A\|^2}{M_{S(n)}\|A\|},$$

тогава

$$\left\|X-\tilde{X}\right\| \le \frac{\left\|\tilde{A}\right\|\left\|F\left(\tilde{X}\right)\right\|+\|A\|\|F(X)\|}{1-M_{S(n)}\|A\|\left\|\tilde{A}\right\|}\|\Delta A\|+\frac{\|\Delta Q\|}{1-M_{S(n)}\|A\|\left\|\tilde{A}\right\|} \equiv pr4.1. \tag{4}$$

Получена е и втора оценка (Remark 4.3 [4])

$$\left\|X-\tilde{X}\right\| \le \frac{\left\|F\left(\tilde{X}\right)\right\|\left(2\|A\|+\|\Delta A\|\right)}{1-M_{S(n)}\|A\|^2}\|\Delta A\|+\frac{\|\Delta Q\|}{1-M_{S(n)}\|A\|^2} \equiv rm4.3. \tag{5}$$

2.1. За уравнението $X + A^* X^{-1} A = Q$

Нека разгледаме пертурбационното уравнение на матричното уравнение (2):

$$\tilde{X} + \tilde{A}^* \tilde{X}^{-1} \tilde{A} = \tilde{Q}, \tag{6}$$

където $\tilde{A} = A + \Delta A$, $\tilde{Q} = Q + \Delta Q$ и $\tilde{X} = X + \Delta X$. Тук ΔA и ΔQ са малки пертурбации на коефициентите A и Q на (2).

Xu [4] изказва следната теорема за оценка на $\frac{\|\Delta X_L\|}{\|X_L\|}$:

Теорема 1. *(Theorem 3.1 [4]) Нека са дадени матриците $A, \tilde{A}, Q, \tilde{Q} \in C^{n\times n}$, като Q и $\tilde{Q}$ са ермитови положително определени. Ако*

$$\|A\|\|Q\|^{-1} < \frac{1}{2}, \quad \left\|\tilde{A}-A\right\| < \frac{1}{2}\left(\frac{1}{2}-\|A\|\left\|Q^{-1}\right\|\right)\left\|Q^{-1}\right\|^{-1}, \quad \left\|\tilde{Q}-Q\right\| \le \left(\frac{1}{2}-\|A\|\left\|Q^{-1}\right\|\right)\left\|Q^{-1}\right\|^{-1},$$

тогава максималните решения X_L и $\tilde{X}_L$ на матричните уравнения (2) *и* (6) *съществуват и удовлетворяват*

$$\frac{\left\|\tilde{X}_L - X_L\right\|}{\|X_L\|} \le \frac{1}{\frac{1}{2}-\|A\|\left\|Q^{-1}\right\|}\left(\frac{\left\|\tilde{A}-A\right\|}{\|A\|}+\frac{\left\|\tilde{Q}-Q\right\|}{\|Q\|}\right) \equiv Th_{XU} \tag{7}$$

Следват някои оценки получени от Hasanov, Ivanov и Uhlig [3].

Теорема 3. *(Theorem 1 [3]) Нека X_L и $\tilde{X}_L$ са максималните положително определени решения, съответно на матричните уравнения (2) и (6) и $\tilde{Q} = \tilde{L}^* \tilde{L}$ е разлагане по Холески. Ако*

$$\left\|\tilde{L}^{-*}\tilde{A}\tilde{L}^{-1}\right\| \le \frac{1}{2} \quad и \quad \tilde{\zeta} = 1 - 2\|A\|\left\|\tilde{Q}^{-1}\right\|\left\|X_L^{-1}A\right\| > 0,$$

тогава

$$\frac{\|\Delta X_L\|}{\|X_L\|} \le \frac{1}{\tilde{\zeta}}\left[\frac{\|\Delta Q\|}{\|Q\|}\frac{\|Q\|}{\|X_L\|}+\frac{\|\Delta A\|}{\|A\|}\frac{2\|A\|}{\|X_L\|}\left\|\tilde{Q}^{-1}\right\|\left(\|A\|+\left\|\tilde{A}\right\|\right)\right] \equiv \tilde{\zeta}_{err}. \tag{8}$$

Теорема 4. *(Theorem 2 [3]) Нека X_L и $\tilde{X}_L$ са максималните положително определени решения, съответно на уравненията (2) и (6). Ако*

$$\left\|\widetilde{A}\right\|\left\|\widetilde{Q}^{-1}\right\| \le \frac{1}{2} \quad и \quad \zeta = 1 - \left\|X^{-1}A\right\| > 0,$$

тогава

$$\frac{\|\Delta X_L\|}{\|X_L\|} \le \frac{1}{\zeta}\left[\frac{\|\Delta Q\|}{\|Q\|}\frac{\|Q\|}{\|X_L\|} + \frac{\|\Delta A\|}{\|A\|}\frac{2\|A\|}{\|X_L\|}\right] \equiv \zeta_{err} \tag{9}$$

Последната разгледана оценка е на Sun и Xu [6]. За да изкажем тяхната оценката разглеждаме операторите:

$\mathbf{L}W = W - B^*WB$, $W \in H^{n\times n}$ ($H^{n\times n}$ - множество на $n\times n$ ермитови матрици),

$\mathbf{P}Z = \mathbf{L}^{-1}(B^*Z + ZB)$, $Z \in C^{n\times n}$ ($C^{n\times n}$ - множество на $n\times n$ комплексни матрици),

където $B = X_L^{-1}A$

$$\alpha = \|A\|_2,\ \beta = \|B\|_2,\ \zeta = \left\|X_L^{-1}\right\|_2,\ p = \|\mathbf{P}\|_U,\ l = \left\|\mathbf{L}^{-1}\right\|_U^{-1},$$

$$\varepsilon = \frac{1}{l}\|\Delta Q\|_U + p\|\Delta A\|_U + \frac{\zeta}{l}\|\Delta A\|_U^2,\ \delta = \frac{\zeta}{l}\left[\left(\alpha + \|\Delta A\|_U\right)\zeta + \beta\right]\|\Delta A\|_U.$$

Теорема 5. *(Theorem 2.1 [6]) Ако*

$$\delta < \min\left\{1,\ \frac{(1-\beta)(\alpha\zeta+\beta)}{l}\right\} \quad и$$

$$\varepsilon < \min\left\{\frac{l(1-\delta)^2}{\zeta\left[l + 2\beta^2 + l\delta + 2\sqrt{(l\delta+\beta^2)(l+\beta^2)}\right]},\ \frac{(1-\delta)\left[(1-\beta)(\alpha\zeta+\beta) - l\delta\right]}{\zeta\left[(1+\beta)(\alpha\zeta+\beta) + l\delta\right]}\right\},$$

тогава пертурбационното уравнение (6) има максимално положително определено решение $\widetilde{X}_L$ *и*

$$\left\|\widetilde{X}_L - X_L\right\|_U \le \frac{2l\varepsilon}{l(1+\zeta\varepsilon+\delta) + \sqrt{l^2(1+\zeta\varepsilon-\delta)^2 - 4\zeta l\varepsilon(l+\beta^2)}} \equiv \xi_* \tag{10}$$

2.2. За уравнение $X - A^*X^{-1}A = Q$

Sun [5] прави пертурбационен анализ на уравнението (1). Оценките получени за това уравнение при определена зависимост между коефициентите може да се прилагат и за уравнението (3). Поради дългото описание на теоремата на Sun [5], тук ще дадем само крайния резултат. Нека X е положително определено решение на (1) и $\widetilde{X}$ е решение на съответното пертурбационно уравнение на (1). При определени условия върху коефициентите на (1) имаме

$$\left\|\widetilde{X} - X\right\|_U \le \frac{2l\varepsilon'}{l - \eta + 1\hat{g}\,\varepsilon' + \sqrt{\left(l - \eta + 1\hat{g}\,\varepsilon'\right)^2 - 4l\hat{g}\left(l - \eta + \hat{\alpha}^2\right)\varepsilon'}} \equiv \widetilde{\xi}_*. \tag{11}$$

За да представим другите резултати разглеждаме пертурбационното уравнение

$$\widetilde{X} - \widetilde{A}^*\widetilde{X}^{-1}\widetilde{A} = \widetilde{Q}. \tag{12}$$

В [3] са дадени следните оценки:

Теорема 6. *(Theorem 3 [3]) Нека* $A, \widetilde{A}, Q, \widetilde{Q} \in C^{n\times n}$, *където* Q *и* $\widetilde{Q}$ *са положително определени. Ако* $\widetilde{\varepsilon} = 1 - \left\|A\right\|\left\|\widetilde{Q}^{-1}\right\|\left\|X^{-1}A\right\| > 0$, *тогава за положително определените решения* X *и* $\widetilde{X}$ *съответно на матричните уравнения* (3) *и* (12) *е изпълнено*

$$\frac{\|\Delta X\|}{\|X\|} \le \frac{1}{\tilde{\varepsilon}}\left[\frac{\|\Delta Q\|}{\|Q\|}\frac{\|Q\|}{\|X\|} + \frac{\|\Delta A\|}{\|A\|}\frac{\|A\|^2}{\|X\|}\|\tilde{Q}^{-1}\|\left(2+\frac{\|\Delta A\|}{\|A\|}\right)\right] \equiv \tilde{\varepsilon}_{err}. \tag{13}$$

Теорема 7. *(Theorem 6 [3]) Нека* $A, \tilde{A}, Q, \tilde{Q} \in C^{n\times n}$, *където* Q *и* $\tilde{Q}$ *са положително определени. Нека*

$$b = 1 - \|X^{-1}A\|^2 + \|X^{-1}\|\|\Delta Q\|, \qquad c = \|\Delta Q\| + 2\|X^{-1}A\|\|\Delta A\| + \|X^{-1}\|\|\Delta A\|^2.$$

Ако

$$\tilde{\varepsilon} = 1 - \|A\|\|\tilde{Q}^{-1}\|\|X^{-1}A\| > 0,\ 1 > \|X^{-1}A\|,\ D = b^2 - 4c\|X^{-1}\| \ge 0,\ \tilde{\varepsilon}_{err} < \min\left\{\frac{1}{\|X^{-1}\|}, \frac{b+\sqrt{D}}{2\|X^{-1}\|}\right\},$$

където $\tilde{\varepsilon}_{err} = \frac{1}{\tilde{\varepsilon}}\left[\|\Delta Q\| + \|\Delta A\|\|\tilde{Q}^{-1}\|\left(2\|A\| + \|\Delta A\|\right)\right]$, *тогава за положително определените решения* X *и* $\tilde{X}$ *на уравненията* (3) *и* (12) *е изпълнено*

$$\|\Delta X\| \le \frac{b-\sqrt{D}}{2\|X^{-1}\|} \equiv S_{err}. \tag{14}$$

3. Числени експерименти

Направени са числени експерименти за двете матрични уравнения с различни матрици-коефициенти и различна размерност на коефициентите.

3.1. За уравнението $X + A^* X^{-1} A = Q$

Аналогично на оценката (14) за уравнение (3), то подобна оценка е вярна и за уравнението (2):

$$S_{err}^{+} = \frac{b-\sqrt{D}}{2\|X_L^{-1}\|}, \quad D = b^2 - 4c\|X_L^{-1}\| \ge 0, \qquad \begin{aligned} b &= 1 - \|X_L^{-1}A\|^2 + \|X_L^{-1}\|\|\Delta Q\| \\ c &= \|\Delta Q\| + 2\|X_L^{-1}A\|\|\Delta A\| + \|X_L^{-1}\|\|\Delta A\|^2. \end{aligned}$$

Отбелязваме, че в оценките на Ran и Reurings $M_{S(n)} = 4\|Q^{-1}\|^2$**, когато** $F(X) = X^{-1}$ **и** $1 - 2\|A\|\|Q^{-1}\| > 0$**.**

Пример 1. Дадено е матричното уравнение $X + A_k X^{-1} A_k = I$ с $A_k = \frac{\delta_k}{\|A\|}A$, където $\delta_k = \frac{1}{2} - 10^{-k}$,

$A = \begin{pmatrix} 4 & 2 & 0 & 0 & 1 \\ 2 & 4 & 2 & 0 & 0 \\ 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 2 & 4 & 2 \\ 1 & 0 & 0 & 2 & 4 \end{pmatrix}$. Имаме матрица $C = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{pmatrix}$ и $A_0 = \frac{C^* + C}{\|C^* + C\|}$

Максималното решение $X_L = \dfrac{I + \sqrt{I - 4A_k^* A_k}}{2}$. Приближението $\tilde{A} = A_k + 10^{-kj} A_0$, а максималното решение на $\tilde{X} + \tilde{A}^* \tilde{X}^{-1} \tilde{A} = \tilde{Q}$ е $\tilde{X}_L = \dfrac{I + \sqrt{I - 4\tilde{A}^* \tilde{A}}}{2}$.

Таблица 1.

j	2	3	4	5
$\dfrac{\|\Delta X_L\|}{\|X_L\|}$	8.65e–06	8.65e–09	8.65e–12	8.13e–15
(4) $\dfrac{pr4.1}{\|X_L\|}$	4.71e–04	4.71e–07	4.71e–10	4.72e–13
(5) $\dfrac{rm4.3}{\|X_L\|}$	4.71e–04	4.71e–07	4.71e–10	4.72e–13
(7) Th_{XU}	2.00e–03	2.00e–06	2.00e–09	2.01e–12
(8) $\tilde{\zeta}_{err}$	3.16e–05	3.17e–08	3.17e–11	3.17e–14
(9) ζ_{err}	3.27e–05	3.27e–08	3.27e–11	3.28e–14
$\dfrac{S_{err}^+}{\|X_L\|}$	1.58e–05	1.59e–08	1.58e–11	1.59e–14
(10) $\dfrac{\xi_*}{\|X_L\|}$	1.58e–05	1.58e–08	1.58e–11	1.59e–14

Пример 2. Разглеждаме матричното уравнение $X + A^* X^{-1} A = Q$, при зададени матрици $X = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ и $A = \dfrac{1}{6}\begin{pmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & 0 & 1 & -1 & -1 \\ 1 & 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$. $C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 0 & 2 & 3 & 3 & 3 \\ 0 & 0 & 3 & 4 & 4 \\ 0 & 0 & 0 & 4 & 5 \end{pmatrix}$, $A_0 = \dfrac{C}{\|C\|}$, $Q_0 = \dfrac{C^* + C}{\|C^* + C\|}$. Нека $\tilde{A} = A + 10^{-2j} A_0$, $\tilde{Q} = Q + 10^{-4j} Q_0$. Решението $\tilde{X}$ е пресметнато с функцията ***dare*** на Matlab ($\tilde{X} = dare(O, I, Q, O, A^*, I)$, където O е нулевата матрица, I е единичната матрица).

Таблица 2.

j	2	3	4	5
$\dfrac{\|\Delta X_L\|}{\|X_L\|}$	7.75e–06	7.75e–09	7.75e–12	7.75e–15

(4)	$\frac{pr4.1}{\|X_L\|}$	2.56e–04	2.58e–07	2.58e–10	2.58e–13
(5)	$\frac{rm4.3}{\|X_L\|}$	2.58e–04	2.57e–07	2.58e–10	2.58e–13
(7)	Th_{XU}	1.68e–03	1.68e–06	1.68e–09	1.68e–12
(8)	$\tilde{\zeta}_{err}$	1.09e–05	1.09e–08	1.09e–11	1.09e–14
(9)	ζ_{err}	1.11e–05	1.11e–08	1.11e–11	1.11e–14
	$\frac{S^{+}_{err}}{\|X_L\|}$	3.15e–05	3.15e–08	3.15e–11	3.15e–14
(10)	$\frac{\xi_*}{\|X_L\|}$	2.70e–05	2.70e–08	2.70e–11	2.70e–14

3.2. За уравнението $X - A^* X^{-1} A = Q$

В оценките на Ran и Reurings $M_{S(n)} = \left\|Q^{-1}\right\|^2$, когато $F(X) = -X^{-1}$.

Пример 3. Разглеждаме матричното уравнение $X - A_k^* X^{-1} A_k = I$ и пертурбационното уравнение $\tilde{X} - \tilde{A}_k^* \tilde{X}^{-1} \tilde{A}_k = I$, където $A_k = \frac{2\delta_k}{\|A\|} A$, $\delta_k = \frac{1}{2} - 10^{-k}$, $A = \begin{pmatrix} 4 & 2 & 0 & 0 & 1 \\ 2 & 4 & 2 & 0 & 0 \\ 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 2 & 4 & 2 \\ 1 & 0 & 0 & 2 & 4 \end{pmatrix}$.

Решенията $X = \frac{I + \sqrt{I + 4A_k^* A_k}}{2}$ и $\tilde{X} = \frac{I + \sqrt{I + 4\tilde{A}_k^* \tilde{A}_k}}{2}$, където $\tilde{A} = A_k + \delta_k^{2^j} A_0$ и $A_0 = \frac{C^* + C}{\|C^* + C\|}$ и C е произволна матрица генерирана с функция ***randn*** на Matlab.

Таблица 3.

j	2	3	4	5
$\frac{\|\Delta X\|}{\|X\|}$	2.21e–02	1.36e–03	5.24e–06	7.75e–11
(4) $\frac{pr4.1}{\|X\|}$	*	2.80e+00	4.55e–03	6.71e–08
(5) $\frac{rm4.3}{\|X\|}$	*	*	4.55e–03	6.70e–08
(11) $\frac{\tilde{\xi}_*}{\|X\|}$	*	*	*	3.07e–07

(13) $\widetilde{\varepsilon}_{err}$	2.06e–01	1.24e–02	4.76e–05	7.03e–10
(14) $\frac{S_{err}}{\|X\|}$	1.12e–01	4.75e–03	1.80e–05	2.66e–10

Звездичките в таблицата означават, че някое условие на теоремата даваща съответната оценка.

Пример 4. Разглеждаме матричното уравнение $X - A^* X^{-1} A = Q$ с

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad Y = \begin{pmatrix} -1 & -1 & 0 & 0 & 0 \\ -1 & 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{pmatrix},$$

$A = \frac{1}{5}\left(Y + \frac{1}{5} I\right)$, $A_0 = \frac{C}{\|C\|}$ и $Q_0 = \frac{C^* + C}{\|C^* + C\|}$, $\widetilde{A} = A + 10^{-2j} A_0$, $\widetilde{Q} = Q + 10^{-4j} Q_0$. Решението $\widetilde{X}$ пресметнато с функцията ***dare*** ($\widetilde{X} = dare(S, I, \widetilde{Q}, R)$, където $S = \widetilde{A}^{-*} \widetilde{A}$ и $R = \widetilde{A} \widetilde{Q}^{-1} \widetilde{A}^*$)

Таблица 4.

j	2	3	4	5
$\frac{\|\Delta X\|}{\|X\|}$	6.70e–06	6.70e–08	6.70e–10	6.70e–12
(4) $\frac{pr4.1}{\|X\|}$	4.71e–05	4.71e–07	4.71e–09	4.71e–11
(5) $\frac{rm4.3}{\|X\|}$	4.71e–05	4.71e–07	4.71e–09	4.71e–11
(11) $\frac{\widetilde{\xi}_*}{\|X\|}$	*	3.34e–03	3.19e–05	3.19e–07
(13) $\widetilde{\varepsilon}_{err}$	4.24e–05	4.24e–07	4.25e–09	4.25e–11
(14) $\frac{S_{err}}{\|X\|}$	2.25e–05	2.25e–07	2.25e–09	2.25e–11

ЛИТЕРАТУРА

1. J. **Engwerda**, On the existence of a positive definite solution of the matrix equation $X + A^T X^{-1} A = I$, *Linear Algebra Appl.,* 194 (1993), 91-108.

2. A. **Ferrante**, B. Levy, Hermitian solution of the equation $X = Q + NX^{-1}N^*$, *Linear Algebra Appl.*, 247 (1996), 359-373.

3. V.I. **Hasanov**, I.G. Ivanov, and F. Uhlig, Improved perturbation estimates for the matrix equations $X \pm A^* X^{-1} A = Q$, *Linear Algebra Appl., /to appear/.*

4. A.C.M. **Ran**, M.C.B. Reurings, On the matrix equation $X + A^* F(X) A = Q$: solution and perturbation theory, *Linear Algebra Appl.*, 346 (2002), 15-26.
5. J.-G. **Sun**, Perturbation theory for algebraic Riccati equations, *SIAM J. Matrix Anal. Appl.*, 19 (1998), 39-65.
6. J.-G. **Sun**, S.-F. Xu, Perturbation analysis of the maximal solution of the matrix equation $X + A^* X^{-1} A = P$ II, *Linear Algebra Appl.* 362 (2003), 211-228.
7. S. F. **Xu**, Perturbation analysis of the maximal solution of the matrix equation $X + A^* X^{-1} A = P$, *Linear Algebra Appl.*, 336 (2001), 61-70.

ГЕОМЕТРИЧЕН ПОДХОД ПРИ ИЗВЕЖДАНЕ НА ФОРМУЛАТА ЗА БРОЯ НА КОМБИНАЦИИ БЕЗ ПОВТОРЕНИЯ

НАТАЛИЯ ХР. ТОНЧЕВА

GEOMETRY APPROACH OF WORKING OUT THE FORMULA FOR THE NUMBER OF THE COMBINATIONS WITHOUT REPETITION

NATALIA HR. TONCHEVA

A different method for introduction to combinatorics is shown in this paper. It is offered a geometry approach of working out the formula for the number of the combinations without repetition.

KEY WORDS: combination, geometry method

1. ВЪВЕДЕНИЕ

В училищния курс по математика комбинаториката, статистиката и вероятностите са твърде изолирани от останалия материал. Примерите и задачите, предложени в повечето учебници отразяват добре приложенията на комбинаториката, но обогатяването им с геометрични примери би допринесло за по-плавен преход от предходните теми (планиметрия) към комбинаторика. От друга страна чрез тези примери допълнително ще се затвърди изученото по планиметрия. Ще се разкрият по-добре приложенията и връзките между отделните теми в учебниците по математика.

В работата се предлага един нетрадиционен подход в реда на въвеждане на съединенията без повторение. Във всички учебници по математика за десети клас се следва схемата на въвеждане – пермутация, вариация, комбинация. Този подход е естествен и се базира на прехода от "простото" към "сложното". Без да отричаме този подход, ще предложим една негова алтернатива – пермутация, комбинация, вариация.

2. ВЪВЕЖДАНЕ НА ПОНЯТИЯТА ПЕРМУТАЦИЯ, КОМБИНАЦИЯ, ВАРИАЦИЯ. ДОКАЗВАНЕ НА НА ФОРМУЛАТА ЗА БРОЯ НА КОМБИНАЦИИ БЕЗ ПОВТОРЕНИЯ

Въвеждането на понятието пермутация и извеждането на формулата за броя на пермутациите предлагаме да следва традиционната методика на преподаване.

Въвеждането на понятието комбинация предлагаме да се осъществи по традиционния начин. Предлагаме да се предложат примерни множества от различно естество – числови, геометрични, предметни и т.н. Да се наблегне на факта, че видът на елементите не влияе на броя на комбинациите без повторения, след което да се пристъпи към извеждането на формулата за броя на комбинациите без повторение на **n** елемента **k**-ти клас.

Основното предимство на този подход е конкретно индуктивното извеждане на формулата за броя на комбинациите без повторение на **n** елемента **k**-ти клас. Фактор за избора на този подход са и нагледността, плавния преход между темите и затвърждаването на предходните знания.

Използваме следната познавателна задача:

Даден е правилен **n**-ъгълник с върхове $\mathbf{A_1, A_2,..., A_n}$. С $\mathbf{m_i}$ означаваме броя на отсечките, свързващи върха $\mathbf{A_i}$ с останалите върхове на **n**-ъгълника.

- $\mathbf{m_1 = ? \; m_i = ?}$ / Отговор **n-1.**
- Колко са отсечките с краища върховете на **n**-ъгълника? / Отговор C_n^2. Извода е на базата на определението.
- Каква е връзката между $\mathbf{m_1 + m_2 + ...+ m_n}$ и броя на отсечките с краища върховете на **n**-ъгълника? Защо? $\mathbf{m_1 + m_2 + ...+ m_n =?}$

От получения резултат стигаме до извода:

$$2C_n^2 = \sum_{i=1}^{n} m_i \qquad (2.1)$$

което води до откриването на интересуващата ни формула:

$$C_n^2 = \frac{n(n-1)}{2} \qquad (2.2)$$

Използвайки същата идея (избор на всевъзможните **k**-ъгълници ($k \le n$), чиито върхове са върхове и на **n**-ъгълника) стигаме до рекурентната зависимост:

$$kC_n^k = nC_{n-1}^{k-1} \qquad (2.3)$$

Прилагайки **к-1** пъти тази зависимост получаваме:

$$C_n^k = \frac{n(n-1)...(n-k+2)(n-k+1)}{k(k-1)...2.1} \qquad (2.4)$$

За въвеждането на понятието вариация ползваме традиционните примери и похвати. Като подчертаваме наредеността на избраните елементи, показваме връзката между броя на комбинациите без повторение на **n** елемента **k**-ти клас и броя на вариациите без повторение на **n** елемента **k**-ти клас, откъдето получаваме формулата:

$$V_n^k = n(n-1)...(n-k+1) \qquad (2.5)$$

3. ЗАТВЪРЖДАВАНЕ

В системата от задачи за затвърждаване е добре да присъстват и задачи ползващи геометрични множества. Това ще допренесе за съзнателното усвояване на материала. При решаването на подобни задачи учениците ще асоциират комбинациите с отсечки, триъгълници, многоъгълници.Тези понятия са им добре познати и ще им е по-лесно да работят с тях, отколото с множества от две, три и повече точки. Такива задачи могат да се намерят в [1] и [2]. Тук ще предложим две задачи.

Задача 1.

Даден е правилен **n**-ъгълник. Да се намери броя на диагоналите. Да се намери броя на триъгълниците, които имат:

1. две общи страни с **n**-ъгълника /Отговор **n**
2. една обща страна с **n**-ъгълника /Отговор **n(n-4).**

Задача 2.

Да се докаже, че броя на диагоналите на правилен **n**-ъгълник е два пъти по-малък от броя на триъгълниците с върхове – върховете на **n**-ъгълника и с поне една обща страна с **n**-ъгълника.

Забележка: В статията работим с правилни **n**-ъгълници с цел да облекчим учениците. По преценка на учителя това условие би могло да се замени с изпъкнал **n**-ъгълник.

4. ЗАКЛЮЧЕНИЕ

В предложения подход няма радикална промяна на методиката на преподаване на съединенията без повторения, а се акцентува на плавния преход между темите, съзнателното усвояване на формулата за броя на комбинациите без повторение на **n** елемента **k**-ти клас, чрез използване на геометричен модел и използване на повече геометрични примери при въвеждане на понятията.

ЛИТЕРАТУРА

4. **Портев** и колектив, *Алгебра*, Летера, 2003.

5. **Портев** и колектив, *20 примерни теми за матура с решения*, Летера., 2003.

ДИСКРЕТНИ СЛУЧАЙНИ ВЕЛИЧИНИ И БАЗОВИ СЪБИТИЯ

НАТАЛИЯ ХР. ТОНЧЕВА, ХРИСТО В. ВЪЛЧЕВ

DISCRET RANDOM VARIABLE AND BASIC EVENTS

NATALIA HR. TONCHEVA, HRISTO V. VALCHEV

In the paper is shown a different method for solving probabilities problems by using basic events. It is shown a method which makes the solution much easier. It is appropriate for students from extracurricular form of studying.

KEY WORDS: basic events, probability, method.

1. Въведение

В тази работа се предлага използването на подходящо избрана пълна група събития, при намирането на редовете (законите) на разпределение на дискретни случайни величини. Този метод облекчава значително решаването на някои задачи.

Определение: Ще казваме, че $H_1, H_2, ..., H_k, ...$ е пълна група събития(ПГС), ако в резултат на провеждания опит сигурно ще се сбъдне точно едно от тези събития.

Определение: Ще казваме, че A е благоприятстващо за B събитие $(A \Rightarrow B)$, ако от сбъдването на A следва сигурното сбъдване на B.

Определение: Ще казваме, че пълната група събития $H_1, H_2, ..., H_k, ...$ е базова за A, ако за всяко H_k е изпълнено едно от двете твърдения: $H_k \Rightarrow A$ или $H_k \Rightarrow \overline{A}$. Събитията $H_1, H_2, ..., H_k, ...$, ще наричаме базови за събитието A.

Определение: Ще казваме, че ПГС $H_1, H_2, ..., H_k, ...$ е базова за дискретната случайна величина ξ, ако тя е базова за всяко едно от събитията $(\xi = x_1)$, $(\xi = x_2)$,..., $(\xi = x_i)$,..., където $x_1, x_2, ..., x_i, ...$ са възможните значения на ξ. Събитията $H_1, H_2, ..., H_k, ...$ ще наричаме базови за дискретната случайна величина ξ.

Ако множеството на елементарните събития е крайно- или безкрайно-изброимо, то групата на елементарните събития е базова, за всяка дискретна случайна величина. Една често използвана базова за ξ група е $(\xi = x_1)$, $(\xi = x_2)$,..., $(\xi = x_i)$,..., където $x_1, x_2, ..., x_i, ...$ са възможните значения на ξ. При решаването на някои задачи, нито една от тези базови групи не е подходяща за намирането на реда на разпределение на дискретни случайни величини.

От дефиницията за базови събития за ξ се вижда, че всяко едно от тях еднозначно определя ξ. Нека означим възможното значение, което ξ ще приеме, ако се сбъдне базовото събитие H_k с $\xi(H_k)$.

2. Теореми

В [1] са доказани следните твърдения:

Т1 /Класическа формула за вероятност-КФВ/ Ако съществува крайна пълна група равновъзможни базови събития (КПГРС) за A, то $P(A) = \frac{M}{N}$, където M е броят на благоприятстващите A базови събития, а N е броят на всички базови събития.

Т2 За всяка крайна или безкрайна, но изброима, базова за A група, вероятността на A е равна на сумата от вероятностите на благоприятстващите A базови събития.

3. Примерни задачи

Задача 1: В урна има n топки, номерирани с числата от едно до n. Всички топки се изваждат последователно без връщане. С ξ_k $(1 \le k \le n)$ е означен номера на k^{-ma} извадена топка. Намерете реда на разпределение на ξ.

Решение: Подходящата базова група в случая е $(\xi_k = 1)$, $(\xi_k = 2)$,..., $(\xi_k = n)$. Това е КПГРС. От Т1 следва, че $P(\xi_k = i) = \frac{1}{n}$, т.е. редът на разпределение на ξ_k е:

ξ_k	1	2	...	i	...	n
$P(\xi_k = i)$	$\frac{1}{n}$	$\frac{1}{n}$	…	$\frac{1}{n}$	…	$\frac{1}{n}$

Задача 2: В урна има 3 бели и 7 черни топки. По случаен начин, последователно с връщане, четири пъти се вади по една топка. С ξ е означен броят на появяванията на бяла топка. Намерете реда на разпределение на ξ.

Решение: При тази задача елементарните събития са много (10000) и не е удобно да изберем за базова групата на елементарните събития. С $\boldsymbol{б_i}$ и $\boldsymbol{ч_i}$ i=1, 2, 3, 4 ще означим събитията – при $i^{-то}$ вадене ще се появи съответно бяла ($\boldsymbol{б_i}$) или черна ($\boldsymbol{ч_i}$) топка. В следващата таблица ще опишем подходяща група на базовите събития:

	Базови събития $H_к$	P(Hк)	ξ(Hк)
H_1	$б_1б_2б_3б_4$	0,0081	4
H_2	$б_1б_2б_3ч_4$	0,0189	3
H_3	$б_1б_2ч_3б_4$	0,0189	3
H_4	$б_1ч_2б_3б_4$	0,0189	3
H_5	$ч_1б_2б_3б_4$	0,0189	3
H_6	$б_1б_2ч_3ч_4$	0,0441	2
H_7	$б_1ч_2б_3ч_4$	0,0441	2
H_8	$б_1ч_2ч_3б_4$	0,0441	2
H_9	$ч_1б_2б_3ч_4$	0,0441	2
H_{10}	$ч_1б_2ч_3б_4$	0,0441	2
H_{11}	$ч_1ч_2б_3б_4$	0,0441	2
H_{12}	$б_1ч_2ч_3ч_4$	0,1029	1
H_{13}	$ч_1б_2ч_3б_4$	0,1029	1
H_{14}	$ч_1ч_2б_3ч_4$	0,1029	1
H_{15}	$ч_1ч_2ч_3б_4$	0,1029	1
H_{16}	$ч_1ч_2ч_3ч_4$	0,2401	0

Тъй като $\boldsymbol{б_i}$ и $\boldsymbol{ч_i}$ са независими събития, то вероятностите на събитията $H_к$ можем да намерим лесно като използваме приложените теореми. Да разгледаме H_8, а за останалите базови събития може да се разсъждава аналогично

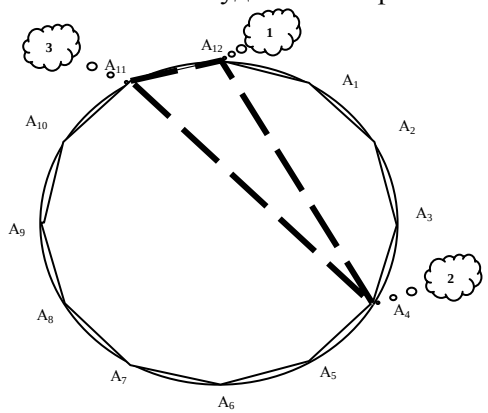
$$P(H_8) = P(б_1 ч_2 ч_3 б_4) = P(б_1)P(ч_2)P(ч_3)P(б_4) = \frac{3}{10}\frac{7}{10}\frac{7}{10}\frac{3}{10} = 0{,}0441$$

Ще разгледаме $(\xi = 2)$. Тъй като благоприятстващите $(\xi = 2)$ базови събития са H_6, H_7, H_8, H_9, H_{10} и H_{11}, то с помощта на **Т2** ще получим, че $P(\xi = 2) = P(H_6) + P(H_7) + P(H_8) + P(H_9) + P(H_{10}) + P(H_{11}) = 6.0{,}0441 = 0{,}2646$ Аналогично се получават $P(\xi = i)$ за $i = 0,1,3,4$. Окончателно:

ξ	0	1	2	3	4
$P(\xi = k)$	0,2401	0,4116	0,2646	0,0756	0,0081

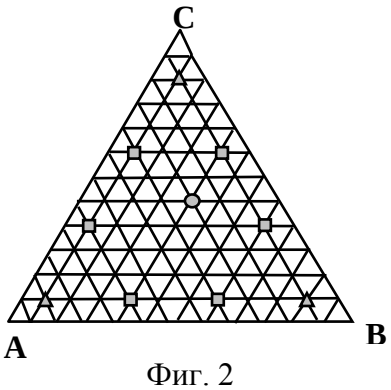
Задача 3: Последователно по случаен начин се избират три от върховете на правилен 12-ъгълник. С Θ_1, Θ_2 и Θ_3 са означени ъглите (в градуси) на триъгълника определен от избраните върхове (номерацията е по реда на избор на връх). Намерете реда на разпределение на Θ_2.

Решение: При тази задача елементарните събития са 12.11.10 = 1320, което прави избора им за базови събития неудобен! На фиг.1 е описано едно от тези 1320 елементарни събития. Означени са избраните върхове и реда на техния избор(A_{12} – първи, A_4 – втори и A_{11} – трети). Всеки един от ъглите ще се "опира" на цяло число дъгички. Всяка от тези дъгички е равна на една дванадесета от описаната около 12-ъгълника окръжност. Следователно Θ_k са кратни на 15^O.



Фиг 1

Ще моделираме тази ситуация по следния начин – в равностранен ΔABC с височина 180 ще прекараме, успоредни на страните му прави на разстояния – 15, 30, 45, 60, 75, 90, 105, 129, 135 и 150. Учениците сами могат да докажат, че $27^{-те}$ прави определят 55 вътрешни за ΔABC точки, за които сумата от разстоянията до страните на триъгълника е 180. Всяка от тези точки ще моделира базово събитие, Θ_1, Θ_2 и Θ_3 са равни на разстоянията съответно до AB, BC и AC. На фиг.2 с кръгче е означено базовото събитие $\Theta_1 = 75^\circ$, $\Theta_2 = 45^\circ$ и $\Theta_3 = 60^\circ$. Моделирани по този начин, $55^{-те}$ събития са равновъзможни и можем да приложим КФВ. С нейна помощ получаваме:



Фиг. 2

θ_2	15^O	30^O	45^O	60^O	75^O	90^O	105^O	120^O	135^O	150^O
$P(\theta_2=X_i^O)$	10/55	9/55	8/55	7/55	6/55	5/55	4/55	3/55	2/55	1/55

Задача 4: Условието на предната задача. Търсят се редовете на разпределение на $\Theta_{(1)}$, $\Theta_{(2)}$ и $\Theta_{(3)}$, където $\Theta_{(1)}$, $\Theta_{(2)}$ и $\Theta_{(3)}$ са стойностите (в градуси) на ъглите Θ_1, Θ_2 и Θ_3, подредени във възходящ ред.

Решение: При тази задача също бихме могли да използваме $55^{-те}$ базови събития от горния

пример, но по-удобно е да изберем, посочените в долната таблица базови събития. С помощта на 55-те базови събития, моделирани на фиг.2 и с помощта на КФВ се получава, че $P(w_1)=\frac{3}{55}$, трите благоприятстващи w_1 базови събития са моделирани с триъгълниче на фиг. 2. $P(w_4)=\frac{6}{55}$, а благоприятстващите w_4 базови събития са моделирани с квадратче на фиг. 2. Аналогично се получават вероятностите на останалите събития. От таблицата, с помощта на **T2** получаваме:

	$\theta_{(1)}$	$\theta_{(2)}$	$\theta_{(3)}$	$P(w_k)$
w_1	15^O	15^O	150^O	3/55
w_2	15^O	30^O	135^O	6/55
w_3	15^O	45^O	120^O	6/55
w_4	15^O	60^O	105^O	6/55
w_5	15^O	75^O	90^O	6/55
w_6	30^O	30^O	120^O	3/55
w_7	30^O	45^O	105^O	6/55
w_8	30^O	60^O	90^O	6/55
w_9	30^O	75^O	75^O	3/55
w_{10}	45^O	45^O	90^O	3/55
w_{11}	45^O	60^O	75^O	6/55

$\theta_{(1)}$	15^O	30^O	45^O	60^O
$P(\theta_{(1)}=x_k^O)$	27/55	18/55	9/55	1/55

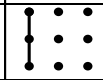
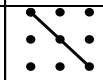
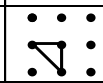
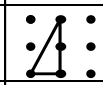
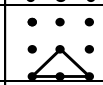
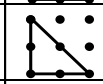
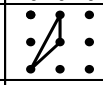
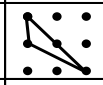
$\theta_{(1)}$	15^O	30^O	45^O	75^O	60^O
$P(\theta_{(2)}=x_k^O)$	3/55	9/55	15/55	19/55	9/55

$\theta_{(1)}$	60^O	75^O	90^O	105^O	120^O	135^O
$P(\theta_{(3)}=x_k^O)$	1/55	9/55	15/55	12/55	9/55	6/55

Задача 5: По случаен начин се избират три от девет точки, които са върховете, средите на страните и пресечната точка на диагоналите на квадрат, с дължина на страната – 4 дм. С S е означено лицето на фигурата, определена от избраните три точки.

Намерете реда на разпределение на:

1) S
2) ξ - дължината на най-дългата отсечка, определена от избраните три точки
3) η - дължината на най-късата отсечка, определена от избраните три точки

	Базови събития H_k	$m(H_k)$	$P(H_k)$	$S(H_k)$	$\xi(H_k)$	$\eta(H_k)$
H_1		6	3/42	0	4	2
H_2		2	1/42	0	$4\sqrt{2}$	$2\sqrt{2}$
H_3		16	4/21	2	$2\sqrt{2}$	2
H_4		16	4/21	4	$2\sqrt{5}$	2
H_5		8	2/21	4	4	$2\sqrt{2}$
H_6		4	1/21	8	$4\sqrt{2}$	4
H_7		4	1/21	8	$2\sqrt{5}$	4
H_8		16	4/21	2	$2\sqrt{5}$	2
H_9		8	2/21	4	$4\sqrt{2}$	2
H_{10}		4	1/21	6	$2\sqrt{5}$	$2\sqrt{2}$

Забележка: три точки на една права определят фигура с нулево лице.

Решение: Базовите събития ще бъдат свързани с конкретна конфигурация на избраните точки. Следващата таблица показва $10^{-те}$ базови събития. Броят на всички конфигурации е $C_9^3=84$. Събитията, определени от тези конфигурации са равновъзможни и ще можем да използваме КФВ.

Базовото събитие H_6 означава, че избраните точки определят равнобедрен правоъгълен триъгълник с катет 4 дм.. Аналогично се определят останалите девет базови събития. С помощта на **T2** ще получим редовете на разпределение на S, ξ и η:

S	0	2	4	6	8
$P(S=x_k)$	2/21	8/21	8/21	1/21	2/21

ξ	$2\sqrt{2}$	4	$2\sqrt{5}$	$4\sqrt{2}$
$P(\xi=y_k)$	8/42	7/42	20/42	7/42

η	2	$2\sqrt{2}$	4
$P(\eta=z_k)$	31/42	7/42	4/42

Въвеждането на базовите за дискретни случайни величини събития, дава възможност на само за просто рашаване на определе клас задачи, но и за съставянето на такива. Например същите 10 базови събития имат случайните величини: най-малкия от ъглите, които сключват трите двойки отсечки, определени от трите точки; най-големия такъв ъгъл и т.н.

Задача 6: Куб, страните на който са оцветени в бял, зелен, червен, син, жълт и оранжев цвят, се разрязва на 27 еднакви кубчета. Последователно, без връщане избираме три от малките кубчета.

Намерете редовете на разпределение на:

1) ξ_i $(i = 0,1,2,3)$ - броят на изваденитe кубчета, i от стените, на които са оцветени;
2) Θ- общият брой оцветени стени, на изваденитe кубчета;
3) æ – броят на изваденитe кубчета с поне една оцветена стена;
4) τ - броят на изваденитe кубчета с поне две оцветени стени;
5) η - броят на изваденитe кубчета, които имат нечетен брой оцветени стени.

Решение: След разрязването на куба, ще се получи едно кубче с неоцветени стени, шест кубчета с една оцветена стена, дванадесет – с по две оцветени стени и осем – с по три оцветени стени. Следващата таблица ще опише базовите събития:

	Базови събития				P(H_k)	Θ(H_k)	æ(H_k)	τ(H_k)	η(H_k)
	ξ_0	ξ_1	ξ_2	ξ_3					
H_1	0	0	0	3	56/2925	9	3	3	3
H_2	0	0	3	0	220/2925	6	3	3	0
H_3	0	3	0	0	20/2925	3	3	0	3
H_4	0	0	1	2	336/2925	8	3	3	2
H_5	0	1	0	2	168/2925	7	3	2	3
H_6	1	0	0	2	28/2925	6	2	2	2
H_7	0	0	2	1	528/2925	7	3	3	1
H_8	0	1	2	0	396/2925	5	3	2	1
H_9	1	0	2	0	66/2925	4	2	2	0
H_{10}	0	2	0	1	120/2925	5	3	1	3
H_{11}	0	2	1	0	180/2925	4	3	1	2
H_{12}	1	2	0	0	15/2925	2	2	0	2
H_{13}	0	1	1	1	576/2925	6	3	2	2
H_{14}	1	0	1	1	96/2925	5	2	2	1
H_{15}	1	1	0	1	48/2925	4	2	1	2
H_{16}	1	1	1	0	72/2925	3	2	1	1

Опитът може да се осъществи по 27.26.25 равновъзможни начина. За да определим вероятностите на H_i, ще се възползваме от "междинна" базова група, свързана с всички възможни комбинации от по три кубчета от общо 27-те кубчета. Тези $C_{27}^3 = 2925$ събития са равновъзможни.

Броят на благоприятстващите H_7 "междинни"събития е $8.C_{12}^2$. По толкова начина могат да се изберат две кубчета с две оцветени стени и едно кубче с една оцветена стена. Прилагайки КФВ ще получим $P(H_7) = \frac{528}{2925}$.

Аналогично се пресмятат вероятностите на останалите базови събития.

С помощта на **T2** лесно се намират търсените редове на разпределение:

θ	2	3	4	5	6	7	8	9
	15/2925	92/2925	294/2925	612/2925	824/2925	696/2925	336/2925	56/2925

æ	2	3
	3/117	104/117

τ	0	1	2	3
	7/585	84/585	266/585	228/585

η	0	1	2	3
	286/2925	1092/1925	1183/2925	364/2925

Учителите сами биха могли да съставят задачи с необходимата им сложност, при които да се прилагат предложените в тази статия методи. Широк кръг от задачи, публикувани в различни сборници по математика ([3], [4], [5]), биха могли да се решат по предложения начин.

ЛИТЕРАТУРА

1. **Вълчев** Хр.В., *Мястото на класическата формула за вероятност в училищния курс по математика*, Сборник на 33$^{-та}$ пролетна конференция, 306-310, София, (2004).
2. **Додунеков** и колектив. *Математика за 12$^{-ти}$ клас (профилирана подготовка)*. Регалия, София, 2003.
3. **Портев** и колектив. *Алгебра*. Летера, 2003.
4. **Портев** и колектив. *20 примерни теми за матура с решения*. Летера, 2003.
5. **Стоянов** Миразчийски, Игнатов, Танушев. *Ръководство за упражнения по теория на вероятностите*. Наука и изкуство, София, 1976.

SITUATIONAL AWARENESS AND INFORMATION INTEROPERABILITY

SANDOR MUNK

Introduction

Knowledge of circumstances and the gathering and processing of information about a situation are essential conditions of every organized activity. This function also plays a significant role in the command and control of military operations. Continuous collection, correlation, analysis, processing, actualization, and dissemination of situational information are done by the staff. Besides this, other information forms the basis for the commander to understand the situation, to forecast the changes in future circumstances, to develop concept of operations, to analyze courses of action, and to evaluate risks.

The result of the collection, processing, and analysis of this information is a mental picture of the situation: the situational awareness. Situational awareness is not simply a mass of collected (acquired), and synthesized information, but also an organized system extended by goal-oriented analysis, and reasoning. Situational awareness is dynamic in nature, it should be actualized as the situation, or the information requirements change.

1. Basics of situational awareness

In psychology and cognitive sciences, ***situational awareness*** is considered as knowledge created through interaction between an agent and its environment. In this sense, awareness can be simply defined as "knowing what is going on". Awareness is an everyday phenomenon, and its role becomes more noticeable as situations and environments become more dynamic, complex, information demanding, and with higher workload, or risks.

Situational awareness knowledge about the relevant environment – essentially is a set of particular information, a system of facts and beliefs about the existence of, characteristics of, and relations between elements of the situation. Characteristics and relations of situation elements are subject, and activity-dependents: naturally different components are relevant in an emergency relief task, in an armed conflict, or in a peace operation. Systematization and analysis of the situational awareness' components can be done based on the situation elements, and the subjects of situational information.

In a given situation the subject of situational awareness does not usually act alone, so among the situation elements we should distinguish the ***players*** acting consciously (in between the subject too), and the ***environmental objects and circumstances*** as the background of the players' activities. Situational awareness generally includes information about existence in time, position in space, state, activity, and changes of, and relations between players and environmental objects.

Situational awareness is created and maintained from environmental effects, and basic (usually raw) data collected, or acquired about the players, and the environment. From these should be selected, correlated and assessed, integrated and synthesized, information that can be used in reasoning to create situational awareness. Naturally not all effects, and data can, or should be used to generate awareness, so not all incoming data will be processed, and the active collection and acquisition of data is done, taking into consideration given requirements.

Situational awareness – as the previous statement had already suggested – contains not only acquired, selected, and synthesized data, but a complete picture can be generated with filling information gaps, and extended by reasoning, or assumptions. This is true in everyday life also, in the case of individual persons, or groups of persons.

The acquisition and use of basic (raw) data – although it could be possible – is unnecessary. Elements of the situation can be grouped by duration of their existence, and their characteristics and relations by rate of their change. So the elements, with regard to a given activity, can be classified as objects with a short, or long lifetime (existence), and the characteristics and relations can be

stable, and variable. Stable characteristics, and relations of long-time existing objects after having acquired, can be stored, and repeatedly used, whereas data about short-lifetime objects, or about variable characteristics of long-existing objects should be acquired repeatedly.

Situational awareness in most cases includes knowledge about spatial characteristics and relations of players, and environmental objects, and in many cases the spatial position and movement is an essential component of the situation. In these cases the elements can be classified as static, and mobile objects. So the stable spatial characteristics of long-life static objects (e.g. terrain features, hydrological objects, buildings, etc.) – as other stable characteristics – can also be stored for future use. In the case of mobile objects, usually not only the actual, changing spatial position, but also the earlier positions – the track of the object – are essential components of situational awareness.

ଌ

Situational awareness has special importance in competitive, or conflicting situations. Competitive situational awareness of economic players (its comparative advantage in awareness compared to their competitors) is commonly accepted as a key condition of efficient decision-making and an essential component of competitive advantages. Knowledge of consumers, competitors, and the economic environment makes possible – among others – to determine characteristics principally influencing the values of products, and services and to increase the efficiency of production, or service. Knowledge of innovative techniques, and procedures is also an essential component of the situational awareness.

In the case of military command and control the environment of a military organization is traditionally the battlefield, so in military literature, instead of situational awareness we find notions of ***battlefield awareness***, or ***battlespace awareness***[4]. Expression 'battlefield' belongs to the traditional land warfare, while 'battlespace'[5] is the term of the joint warfare of the XXI century. In addition to battlespace awareness, in many publications and professional documents we can find the ***battlespace knowledge*** expression too.[6] In these materials the meaning of this expression corresponds to the referenced meaning of battlespace awareness, and the latter used as a synonym, or with a narrower meaning, indicating a lower, intermediate level of the situational awareness.

In addition to the interpretation of situational awareness as particular knowledge about a situation, there is another interpretation that expresses a general capability to generate the situational awareness. This capability is based on modern information collecting (sensing, observing, etc.), and information processing means, connected in a unified system by different networks. This interpretation appears in the terms ***dominant battlespace awareness***, or dominant battlespace knowledge, that are considered essential capabilities of the XXI. Century warfare in American military visions, and plans.

ଌ

Situational awareness can be interpreted not only in the case of individual agents (persons, etc.), but also in the case of systems of loosely, or tightly cooperating agents (e.g. organizations, groups). In these systems, agents with identical, or different capabilities work together in a coordinated way, to achieve common, or agreed goals (objectives). In consequence of their different capabilities, and activities, these individual agents, as appropriate, have their own unique situational awareness, more or less different from others'. The complex (set, or system) of this awareness constitutes the integrated, shared awareness of the organization or group.

---

[4] ***Battlespace awareness***: awareness of the battlespace yielding an interactive "picture" which provides timely, relevant and accurate assessment of friendly and enemy operations within battlespace. [Concept for Future Joint Operations, Part II – Terms and Definitions, p 83]

[5] ***Battlespace***: The environment, factors, and conditions which must be understood to successfully aply comat power, protect the force, or complete the mission. This includes the air, land, sea, space, and included enemy, and friendly forces, facilities, weather, terrain, the electromagnetic spectrum, and the information environment within the operational areas and areas of interest. [Joint Pub 1-02, p 57]

[6] See for example [WALTZ], or [JOHNSON-LIBICKI].

Shared awareness is not only a summative set of its components, but a coordinated, continuously harmonized system, that in the case of complex organizations, can be even multilevel. Situational awareness of individual agents – according to the level and contents of cooperation between them – can overlap each other: to effectively, and successfully accomplish the related activities, they must have identical, or equivalent knowledge of the affected elements of the situation. In addition to these overlaps, naturally any agent's situational awareness can also contain other pieces of knowledge, typical for, and necessary to, the given agent.

In a cooperating group, or organization, the shared situational awareness, according to the division of labor, can be divided by spatial, or functional characteristics, and by levels. In case of spatial division (e.g. to adjacent operational areas) overlapping parts relate to the adjacent parts. That means the knowledge of the "neighbors' situation". In case of functional division (e.g. to arms, and services) the overlapping part contains objects with relevance for both agents. In both cases, the detail, and precision of the individual agents' situational awareness can, or usually will be, different. Similarly, different detail (precision) characterizes the situational awareness of different command levels in a complex organization.

2. Creation and maintenance of shared situational awareness

The creation and maintenance of situational awareness is accomplished using and synthesizing data coming from information collecting devices, and information gathered from other sources previously, or during execution of operations and coordinating situational awareness of different players. So the application components supporting creation and maintenance of situational awareness include applications processing incoming messages, gathering information directly from different sources, and correlating situational information.

Processing of incoming messages containing situational information include processing of messages of different types, and nature. From the viewpoint of military command and control, and battlefield awareness, the most significant messages containing situational information can be classified into two basic categories: bit-oriented data links, and character-oriented message formats. The original purpose of the former category was to handle (near-) real time information exchange between command posts, and weapons (weapon systems), and their platforms, while the latter were means of information exchange between commands, political or governmental, and other organizations.

Synchronized knowledge of the battlespace situation, and its visualization, are essential for every cooperating organization, and all of their autonomous, functional units. This knowledge forms the basis of the common, coordinated activities of these players. The pieces of common situational knowledge are visualized in the form of different "pictures" usually attributed with the terms 'recognized', and 'common'.

Recognized pictures (air, maritime, land/ground) contain assessed information about essential objects (and their descriptive, and spatial characteristics) in a given area, belonging to a given dimension of the battlespace (air, sea, land, etc.). They usually contain minimal map basics (e.g. land and water bodies, boundaries of countries, hydrography, main populated places, and roads) to help determine the real geographical positions, and items to represent position, and movement of the relevant objects of the situation. In the military application, the representation of these situational (generally mobile) objects is usually accomplished in the form of so-called tracks. Track is a basic concept of military situation visualization that represents the actual spatial position of a mobile battlefield object, directly observed, or determined, using data from different sources.

Common pictures provide a uniform, clear representation of knowledge about all relevant elements (players, and their environment), for a given command level. Common pictures describe the actual situation, and contain components, that help commanders to anticipate and influence future situations. The two basic types of these pictures – depending on the command level, and the area represented – are common operational picture, and common tactical picture.

Common Operational Picture (COP) is the complex of situational information available for a commander with an area of responsibility. In the NATO command structure such commanders are, for example, the regional commanders, or in the case of the U.S. armed forces, the combatant commanders. In those states, where there are no military commanders with a unique area of responsibility, the common operational picture is connected with the level of Joint Staff, and the Chief of Staff. Naturally COPs of different areas of responsibility, or union of them should be made available to the higher (national military, national political, alliance, or coalition) command levels, since they are necessary for those levels to determine military missions appropriate to reach given defense, or security objectives, and to control or supervise the execution of assigned military operations, and activities.

Common Tactical Picture is a concept connected with a given operation, the appropriate area of operations, and with the (multinational, or national) Joint Task Force executing this operation. Common Operational Picture of a given area of responsibility is basically generated, and maintained using the recognized pictures of the component commands of the given area (command), and the common tactical pictures of the operations executed in the given area, completed with information from other sources if required. Similarly recognized, or common tactical pictures can be created from other, or lower levels (e.g. local air picture, subordinated unit's tactical picture, etc.) by integration, coordination, and completion.

The contents of a recognized, or common picture, and the shared situational awareness represented by them, are always the responsibility of the the given commander (chief, etc.) . It is his/her responsibility to interpret, or extend incomplete information, to harmonize conflicting information, and to determine information he/she "owns". So the situational picture of a higher command level determines the situational pictures (situational awareness) of subordinates, because the commander's assessment and concept forms the basis of the given operation and the basic condition of the coordinated execution.

In case of alliance, or coalition operations in general, situational awareness of the following command structure elements must be coordinated as required:

- allied commander responsible for the given area of operation, and the higher (military and political) command levels of the alliance;
- national commands (MODs, Joint Staffs, etc.) of the armed forces participating in the given operation directly, or indirectly;
- commander of the (combined) joint task force executing the given operation;
- component commanders of the (combined) joint task force;
- national contingent commanders of troop contributing nations' forces;
- commanders of larger national units participating in operation.

According to previous definitions, the first two elements have a common operational picture, and the last three have a common tactical picture.

In NATO operations, or other operations mandated by international organizations for security (NATO, OSCE, etc.), and led by NATO, there can appear other elements. For example:

- component commands of NATO regional commands;
- joint sub-regional commands as parent HQs of a combined joint task force HQs;
- and decision-making bodies and organs of other organizations (UN, OSCE, etc.).

As it can be seen from the previous two lists, that in a multinational operation there, are complex relations between situational awareness of different command levels, commands that strongly affect the contents, characteristics, and the amount of information exchange needed to maintain this awareness.

3. Role of information interoperability in maintaining situational awareness

Creation and maintenance of common situational awareness requires continuous and widespread information exchange between cooperating actors. This includes sharing of situational

information coming from different sources, and harmonizing individual situational awarenesses of the actors. During military operations of our age information exchange should be realized between actors (organizations, groupings) characterized by, in some respect, increasing heterogeneity in different areas and to different degrees.

Heterogeneity of forces conducting operations – even in the case of armed forces of a given state – exists in many different areas, and a lot of them are unavoidable. The first and most fundamental type of heterogeneity comes from the differences based on the division of labor, the field specialization that is an inherent characteristic of large organizations, organizational systems. In the case of armed forces, the field specialization appears in the form of armed services, and arms.

The next version of heterogeneity appears in the differences in technical systems, and equipments of organizations, organizational elements with same functions. Units equipped with different technical means, but with identical main functionality have, to some extent different capabilities, their state can be described with partly different information, and in order to fulfill their tasks they require partly different information. Heterogeneity in equipments is mostly natural, and mostly unavoidable. The reasons lie in the specialties of the technological development, and in the requirements of organizational effectiveness, and economy.

Third version of heterogeneity is in the execution of tasks, and in the procedures and methods applied by organizations with the same functionality, and equipment. These differences arise from the variances in the superior environment (doctrines, directives, regulations, etc.), the personal qualities and attitudes of the members of the organization, and the organizational traditions, cultures. This leads to the realization of the emerging significance of national, and cultural differences, and to the questions of combined joint operations.

Nowadays in case of forces conducting military operations, and in consequence of increasing importance of multinational operations, degree and role of heterogeneity forms discussed above have essentially grown. Continually growing heterogeneity – among others in doctrines, training levels, technological levels, or cultures – is one of the basic features of 20th century warfare, and recent military operations. Roots of this process can be found in the third part of the 20th century, and the reasons of its acceleration are the essential changes occurred in the threats, and challenges requiring military responses, and in the security environment.

ꕥ

In a heterogeneous environment the fundamental condition of creation and maintenance of shared situational awareness is the ***information interoperability*** of cooperating forces, that is mutual capability of different actors necessary to ensure exchange and common understanding of information needed for their successful cooperation. In the Information Age, in addition to information interoperability between people, or organizations, increases the role, and importance of information interoperability between people and information systems (applications), and between information systems themselves.

In addition to traditional information systems, in our age continually increases the number of other types of systems, and devices to collect (gather, acquire) information, or to perform different tasks, that have information exchange capabilities. In network centric warfare literature the three groups mentioned before are: command and control support systems, sensor systems, and engagement ("shooter") systems. In the following by information system we mean any system with information exchange capabilities.

Information systems obviously are not able to interpret data stored, or processed, they 'do not know' their meaning. Only people utilizing these devices can assign meaning to data. So only an intended meaning, planned interpretation can be assigned to knowledge components stored on information media, or in information systems (and any devices with information functions), and it should be connected to users of these systems (devices), or to information providers. The intended meaning, agreed intentions, and interpretation of the primary users are usually determined in the purpose and functional requirements of the given system.

So ***interoperability between information systems*** is a mutual capability to exchange of data preserving their intended meaning. Without human assistance this requires the transmission, exchange and, if necessary, transformation of representations between information systems. Increasing role, and importance of information system's interoperability is the consequence of the significant improvements in information exchange capabilities of information systems, applications, first of all the spread and globalization of computer networks.

Information interoperability requires different components, functions, and capabilities, that are almost uniformly classified into three groups (three levels), built on top of each other. The technical level of information interoperability is a complex of capabilities to handle physical (material) representations of information. It is the foundation, a necessary prerequisite of efficient information exchange. The syntactical level of information interoperability includes capabilities to handle intermediary representations based on physical representations, related to languages, message- and data formats used in information exchange. At last the semantical level information interoperability is a group of capabilities connected with the interpretation (meaning) of syntactical level representations, with the exchange of information preserving intended meaning.

Nowadays in information technology the degree of heterogeneity at the first two levels has significantly decreased, and expected to decrease in the near future. Both physical transmission means and message formats have dominated by standardized solutions. So in our days interoperability between cooperating information systems on the two first levels can be realized relatively easily. Much more difficult to ensure in case of different systems the common interpretation (semantic interoperability) of data transmitted in messages, stored in, or accessed from databases, and functions, services provided by information systems.

The role, and importance of information, and semantic interoperability in creation and maintenance of situational awareness was emphasized in a report of the Department of Defense of the United States to the Congress as follows: "Network Centric Warfare is based upon the ability of a force to develop shared situational awareness in the cognitive domain. Technical interoperability will get us to the point where the information is correctly represented in distributed systems, but does not ensure that the individuals in different locations, in different organizations, at different echelons have a similar understanding even though they 'see' the same thing. With the added complexity of coalition operations that involve different cultures, the problem is greatly compounded. Semantic interoperability is the capability to routinely translate the same information into the same understanding. This is, of course, necessary to develop the shared situational awareness upon which mature forms of Network Centric Warfare are based."[7]

Summary

Situational awareness is an inherent feature, an essential condition of existence and activity (operation) of every active, goal-oriented, autonomous object, being in interaction with its environment – knowledge about the relevant environment, including facts and beliefs about existence of, characteristics of, and relations between elements of the situation.

In military practice, situational knowledge is visualized in the form of different "pictures". Recognized pictures contain assessed information about essential objects in a given area, belonging to a given dimension of the battlespace. Common pictures provide a uniform, clear representation of knowledge about all relevant elements (players, and their environment), for a given command level.

Creation and maintenance of shared situational awareness requires efficient information exchange between cooperating forces, characterized by different kinds of heterogeneities. So interoperability of information systems is a necessary prerequisite of shared situational awareness. Nowadays interoperability at physical, and syntactical levels can be realized relatively easily, but the role of semantic interoperability has significantly increased.

[7] Network Centric Warfare, DoD Report to Congress. 27 July 2001. [p 6-2]

REFERENCES

Concept for Future Joint Operations. Expanding Joint Vision 2010. – Joint Chiefs of Staff, May 1997.

Global Command and Control System (GCCS), Common Operational Picture (COP) Handbook for GCCS 3.02, Version 2.0 – Joint Chiefs of Staff, 31 July 1998.

Joint Publication 1-02, DoD Dictionary of Military and Associated Terms. – Joint Chiefs of Staff, 24 June 2000.

JOHNSON, Stuart – LIBICKI, Martin (ed.): *Dominant Battlespace Knowledge.* – National Defense University Press Book, October 1995.

Network Centric Warfare. Department of Defense Report to Congress. – Department of Defense USA, 2001.

WALTZ, Edward: *Information Warfare Principles and Operations.* – Artech House, Boston-London, 1998.

CONCEPTS OF INTEROPERABILITY AND THEIR RELATIONS

SANDOR MUNK

In the professional and popular literature one can find a lot of expressions, definitions, and understandings of the concept of interoperability. Albeit these concepts are obviously strongly connected, there are no widely accepted definitions, and only a few publications discussed the relations between different types of interoperability. This publication analyses the existing definitions of interoperability, and suggests some new general definitions, finally discusses the different types of interoperability, and their relations.

KEYWORDS: interoperability, types of interoperability, organizational interoperability, information interoperability, technical interoperability.

Introduction

From the end of the 20th century, in the emerging Information Age, information and information capabilities possessed by different actors are increasingly significant resources used to effectively and efficiently fulfil their activities. The actors usually don't act alone, they can only achieve their goals in cooperation, coordinated with other actor's. In addition to common goals, the most important elements constituting the foundation of cooperation are infosphere components: common and shared situational awareness, and plans. Moreover these elements require other knowledge components (e.g. concepts with common understanding, jointly accepted principles, and rules). So the goal-oriented use, and operation of cooperating partners' knowledge (information) and information processes require specific abilities, additional to their individual information capabilities.

Role of information capabilities necessary to cooperation is especially significant in execution of those complex operations that require coordinated actions of a mission-oriented and in many cases dynamically changing group of several actors. In military practice this significance has became essential with appearing of the combined joint task force (CJTF) concept, and operations other then war, where military organizations cooperate with a lot of other government, or non-government (NGO, PVO) organizations.

Nowadays, in addition to human factors, the expertise and experience of military leaders and professionals participating in operations, information and other knowledge components embodied in IT applications already form one of the most important components of information capabilities. So to efficiently support military operations, to develop required information capabilities, and to gain and maintain information superiority in conflicts it is necessary to establish a task- and situation-oriented, information exchange among human actors, and IT components extending from exchange of raw data, to exchange of highly synthesized knowledge components.

Cooperation capability of actors in the infosphere, or interoperability is a frequently used, "popular" concept of our days' military - and not only the military – literature, and documents. A lot of publications discussed different aspects of interoperability, but only a few of them tried to use an all-embracing, system-oriented approach. The objective of this publication is to build a coherent conceptual foundation of and a framework for the different concepts of interoperability.

Definitions of interoperability

First appearance of the concept of interoperability is connected to military application. Expression 'interoperability' appears in a dictionary of English language in connection with expression 'interoperable', defined as "capable of being used or operated reciprocally." [1, p 997] In the dictionary, appearance of this expression estimated to 1965-70, and the example used is 'interoperable weapons systems'.

In 1992, based on C4I experiences during Desert Shield/Desert Storm, interoperability has appeared as a basic type of capabilities necessary to efficient, and effective cooperation, in a vision document, "C4I for the Warrior" [2], and has became an important concept of other military visions, doctrinal documents, and directives. The meaning of this concept in military application, first connected mainly with technical aspects, and interoperation between, and among C4I systems, was gradually extended to organizations, groupings, and forces.

This latter appears on security policy level in "Defence Capabilities Initiative", launched by NATO Heads of State and Government, at Washington in 1999. According to the first point "The objective of this initiative is to improve defence capabilities to ensure the effectiveness of future multinational operations across the full spectrum of Alliance missions in the present and foreseeable security environment with a special focus on improving interoperability among Alliance forces, and where applicable also between Alliance and Partner forces." [3, p 61-62] The same has formulated in the basic NATO publication providing 'capstone' doctrine of allied joint operations: "The effectiveness of Allied forces in peace, crisis or in conflict, depends on the ability of the forces provided to operate together effectively and efficiently. … A common doctrine supported by standardisation of equipment and procedures, validated through participation in joint and multinational training exercises, provides the basis for the formations and units of a joint and multinational force to be able to work together." [4]

Concept of interoperability appears in numerous – military, and other – publications, document also, obviously in many cases with different meaning, or emphasis. Definitions are changing even in different versions of the same publication (e.g. glossary, see for example [5], [6], and [7]). Let enumerate the most important definitions to use as a basis for analysis, and attempting to formulate a generalized definition of interoperability.

The ability of Alliance forces and, when appropriate, forces of Partner and other nations to train, exercise and operate effectively together in the execution of assigned missions and tasks. [6, p 2-I-6, after 15/7/2000]

The ability of systems, units or forces to provide services to and accept from other systems, units or forces and to use the services so exchanged to enable them to operate effectively together. [5, p 2-I-7; 8; 9, p 215]

(Communications and Information Systems): Ability to provide services and information to and accept from other systems and to use the services and information so exchanged to enable them to operate effectively together. [8]

The condition achieved among communications-electronics systems or items of communications-electronics equipment when information or services can be exchanged directly and satisfactorily between them and/or their users. The degree of interoperability should be defined when referring to specific cases. [9, p 215]

The capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units. [10, p 2-1-4; 11, p 12]

Finally let see two definitions from the Hungarian military literature, that are rather descriptive, than definitive:

An expression of ability to take part in international cooperation. Regarding to armed forces it supposes, that within a given coalition command and control, and armament of military organizations, expertise of personnel makes cooperation, communication between staffs, connections between (compatibility of) their weapon systems, coordinated system of supply with maps, ammunition, and fuel, etc. that is to say everything needed to reach the assigned objective possible. [12, p 596]

For successful cooperation it is necessary common doctrinal base, mutual knowledge of capabilities of each other, required level of coordination, similar professional thinking, appropriate language practice, unified decision methods, and joint exercises. [13, p 23]

It can be seen from definitions enumerated, that they describe different concepts, although all these concepts have common characteristics, so we can hypothesize, that these are different forms (types) of interoperability in general. The conclusion is, that we need a more general definition of

interoperability, a general concept, that the concepts described in the above definitions are specializations of.

The definitions introduced above show that interoperability has two basic characteristics. First it is a relation, a mutual capability between/among two or more objects. Secondly it is a functional capability strongly connected with, and supporting cooperation. So the suggested definition of interoperability in general is the following:

Interoperability is a relation between/ among objects, a mutual capability necessary to ensure successful and efficient interoperation, supporting cooperation.

To fully understand the meaning of interoperability we should analyse the relation between cooperation, and interoperation. Based on the commonly accepted meaning of the two terms, it should be clearly stated, that they are not synonyms. Cooperation includes existence of a common goal, consciousness, deliberateness, and agreement, but interoperation by all means does not. So on the basis of all these interoperability, as a capability to interoperate, is a necessary, but not sufficient condition of cooperation.

Subjects of interoperability – appearing in definitions – are active objects that can be grouped into two basic types: organised groups of humans acting consciously (forces, groupings, organisations, etc.), or technical systems operating purposefully (equipments, functional units, etc.). Accordingly we can differentiate interoperability between/among actors, and between/among equipments.

Types of interoperability

In the military literature these to types are most frequently called operational interoperability, and technical interoperability. In the non-military area we can find other, similar definitions. For example a European eGovernment initiative [14, p 16] uses the concepts of:

- organisational interoperability, that "is concerned with defining business goals, modelling business processes and bringing about the collaboration of administrations";
- semantic interoperability, that "is concerned with ensuring that the precise meaning of exchanged information is understandable by any other application that was not initially developed for this purpose", and;
- technical interoperability, that covers the technical issues of linking computer systems and services".

The two types (operational/organizational vs. technical interoperability), although they have many similarities, have significant special characteristics as well. From the two, the first plays the primary role, and the technical interoperability is subordinated to him. The later is necessary just to create the conditions for the cooperation between/among actors, because successful and efficient cooperation requires not only interconnection, and harmonization of traditional organizational processes, but also those of implemented, or supported by technical systems.

In consequence of revolutionary improvement of information technology, spreading and growing importance of information activities, functions, and devices, the role, importance, and some characteristics of technical interoperability have changed. In addition to interoperability between devices supporting interoperability, and being resources of particular actors, has appeared an other type of technical interoperability between equipments of an actor, and the systems of the information environment.

In the literature 'interoperability' appears in many expressions with different attributes, that indicate different types, or components of interoperability, being in connection with, based on each other. All of the definitions we listed in previous section describe 'interoperability' without any attribute, but actually they are definitions of different types of interoperability, so we should carefully distinguish between them. In addition to operational/organizational interoperability, semantic, and technical interoperability, there appear expressions such as: command-and-control (C2) interoperability, intelligence interoperability, logistics interoperability, and information interoperability.

Role, and significance of different interoperability types are not the same, so at first we should determine the one that has primacy. It is commonly accepted, that operational/organizational interoperability plays primary role among interoperability types, because the common purpose of them is to support cooperation. This primary concept can be defined as follows:

Operational/organizational interoperability is a relation between/among actors cooperating to achieve a common goal, an overall, mutual capability necessary to ensure successful and efficient cooperation.

Successful and efficient cooperation is founded on continuous harmonization of goals and situational awareness, coordinated planning and execution of common activities, that require systematic information exchange (communications), and a shared ontology (system of concepts) as a base of all these. In the course of these common activities information, material goods, and services flow between the cooperating partners.

The most comprehensive version of interoperability between cooperating actors demands an appropriate level of interoperability in every functional area, because an efficient cooperation is not possible without required level of harmonization between appropriate organizational functions (functional processes). According to this fact, interoperability can be defined regarding any functional area, as a mutual capability to cooperate in this area.

A functional area interoperability is a relation between/among actors cooperating to achieve a common goal, a mutual capability necessary to ensure successful and efficient cooperation in a given functional area.

Different functional area interoperability types have different significance. The primary role by all means belongs to command and control (C2) interoperability, because the C2 processes create and maintain basic conditions of cooperation between actors. The second group contains those functional area interoperability types that are connected with basic organizational processes, specific for the given actor. And finally the third group contains functional area interoperabilities common to all actors, related to general organizational areas (logistics, human resource management, etc.).

Fundamental condition of successful and efficient operation of complex organizations, organizational systems, groupings is the sufficient level of information exchange between components, the sharing, and coordinated exploitation of information necessary for cooperation. In case of heterogeneous components there is an additional condition, the information interoperability of the cooperating parties that in our opinion, in its broader sense can be defined in the following way:

Information interoperability is a mutual capability of different actors necessary to ensure exchange and common understanding of information needed for their successful cooperation.

Information interoperability necessary for information exchange between actors, and utilization of information resources can be divided into different components, into different levels. In the literature it is commonly accepted a three-level system. The first is the physical (material) level of representations, mediums used in information exchange, and information gathering; the second is the syntactical level of languages, message- and data formats; and the third is the semantic level of content, and meaning to exchange.

As follows from the above definition, information interoperability has two basic components: ability to exchange information, and ability to develop a common, shared interpretation of this information. In the professional and popular literature there are other expressions connected with information interoperability, such as: language interoperability, conceptual interoperability, or intellectual interoperability.

REFERENCES

[1] *Webster's Encyclopedic Unabridged Dictionary of the English Language.* Gramercy Books, New York, 1996.

[2] *C4I for the Warrior, "The Joint Vision for C4I Interoperability".* Joint Chiefs of Staff, 1998.

[3] Defence Capabilities Initiative. In. *The Reader's Guide to the NATO Summit in Washington. 23-25 April 1999.* NATO Office of Information and Press, Brussels, 1999.

[4] *AJP-01(B), Allied Joint Doctrine.* Ratification Draft 1. NATO Standardization Agency, 2000.

[5] *AAP-6(V), NATO Glossary of Terms and Definitions (English and French).* NATO Military Agency for Standardization, Brussels, 1998.

[6] *AAP-6(V), NATO Glossary of Terms and Definitions (English and French).* Modified Version 02. NATO Military Agency for Standardization, Brussels, 2000.

[7] *AAP-6(2006), NATO Glossary of Terms and Definitions (English and French).* NATO Standardization Agency, Brussels, 2006.

[8] *AAP-31(A) NATO Glossary of Communication and Information Systems Terms and Definitions.* NATO C3 Agency, 1998.

[9] *Joint Publication 1-02, Department of Defense Dictionary of Military and Associated Terms.* Joint Chiefs of Staff, 2001.

[10] *ADatP-02 (H), NATO Information Technology Glossary.* NATO Military Agency for Standardization, 2000.

[11] *ISO/IEC 2382-1, Information technology – Vocabulary, Part 1: Fundamental Terms.* Third edition. 1993.

[12] *Hadtudományi Lexikon.* Magyar Hadtudományi Társaság, Budapest, 1995.

[13] *Katonai kislexikon.* Honvéd Vezérkar Tudományszervező Osztály, Budapest, 2000.

[14] *European Interoperability Framework for Pan-European eGovernment Services.* Version 1.0. European Communities, 2004.

NECESSITY OF A MILITARY IT INTEROPERABILITY INFRASTRUCTURE

SANDOR MUNK

In our world of globalisation the cooperation, and interoperability between/among different actors in every sphere (political, defence, economical, etc.) plays a more and more important role. In the changing international security environment the traditional interoperability solutions for military IT systems are less and less appropriate. In a dynamically changing environment, to cooperate with increasingly heterogeneous systems, new methods are necessary. An infrastructure-based approach provides a possible solution to ensure information interoperability.

KEYWORDS: military informatics, military information environment, information interoperability, interoperability infrastructure.

Introduction

The growing role of cooperation, and the evolution of Information Age has as a consequence the increasingly growing importance of information interoperability. Nowadays successful and efficient activity, or operation of actors (individuals, organizations, systems) essentially unthinkable without extensive information exchange between actors, and without widespread use of different information sources, information services of the infosphere.

Prior to the NATO Prague Summit NATO defence ministers identified four key operational capability areas. These included the improvements in interoperability of deployed forces. [1, p 27] Operational interoperability requires appropriate level of interoperability on different functional areas (such as command and control, intelligence, logistics, etc.). All functional area interoperabilities are based on information interoperability, and technical interoperability.

An essential condition of ensuring the interoperability types mentioned above is interoperability of military IT systems that in practice first appeared in case of actors working on similar functional areas, and being in a permanent and close cooperation. Traditional IT systems' interoperability solutions, based on standardized protocols and intermediary representations (bit- and character-oriented message formats, common data models), were developed in support of these kind of cooperation. In the changing international security environment, and as a consequence of changes in nature of military operations, and structure of forces, the traditional interoperability solution is less and less appropriate.

Traditional ways, and methods of creating, and maintaining information interoperability are less and less effective on the Information Age battlesphere. In a dynamically changing environment, to cooperate with increasingly heterogeneous systems, new methods are necessary. An infrastructure-based approach provides a possible solution. This publication summarizes the basics of information interoperability, discusses some consequences of network centric warfare regarding information interoperability, and finally outlines a new, network centric approach of an interoperability layer, as a part of the information infrastructure.

Military interoperability environment and its changes

Concept of interoperability and information interoperability is essentially a relational concept. According to the commonly accepted understanding a specific actor, system, device can not be interoperable in itself, but only related to a well-defined group of actors, systems, devices, in cooperation (inter-operation) with them. So information interoperability is a mutual capability of different actors necessary to ensure exchange of common understanding of information needed for their successful cooperation. Information needs of cooperation are expressed in form of information exchange requirements that define: what kind of information need to be exchanged, between what actors, and in what ways (on what carriers, in what quantity, and quality).

Information exchange between/among cooperating actors can be done with the help of intermediary representations. To ensure efficient cooperation, and meaning-preserving information exchange, it is necessary to select, or develop intermediary representations, and to determine an agreed, common understanding of these representations. So interoperability of a specific actor regarding a given community of interest means, that he/she is able to use the intermediary representation(s) of the given community, in other words he/she can send information, messages, questions, or receive information, messages, replies with appropriate contents, and in appropriate formats, as if he/she could speak the "language" used in the given community.

From the point of view of interoperability between actors, information exchange without human assistance between their IT systems is of a continually growing importance. During exchange, and if necessary, transformation of data stored, handled in IT systems, it is necessary to ensure, that source and target data carry the same (or similar enough) meaning for cooperation.

So IT interoperability is a mutual capability of IT systems, devices, and applications to – if necessary after intermediary transformations – receive, exchange data, preserving the meaning assigned to data by the primary user community. In case of IT systems interoperability is usually described not in respect of an explicit community of cooperating actors, but rather in connection with a given intermediary representation.

The beginning of the XXI. century is characterized by significant changes in the international security environment, the nature of military operations, the missions, and structure of military forces executing operations, and in doctrinal principles. These changes have definite influence on requirements of interoperability between military IT systems, and on possible ways, and methods of ensuring interoperability. In the following we will analyze and summarize the most important changes has happened (or will probably happen), and their consequences. For this purpose we will take the common vision of the two NATO strategic commanders as basis. [2]

One of the basic elements of the allied commanders' strategic vision is the holistic approach of military operations, and the extension of their relations to other – informational, economical, social, legal, and diplomatic – activities. [2 (points 11., 13., 18.)] This involves significant extension in, and continuous development, and changes of information used in preparation and execution of operations.

Other significant element of strategic vision is the change in structure of forces executing military operations, the extension of the circle of cooperation partners, and the evolving dominance of multilateralism. Military operations of our age are planned and executed in a joint, combined (allied, and even coalition) framework, usually established for the given mission, and based on occasional national offerings. To fulfill their mission the executing forces must establish close cooperation with other, non-military – international, governmental, non-governmental, and civil – organizations. [2 (points 17., 21., 22., 23.)]

From the changes, and characteristics presented before it follows that a given military organization, and its IT system(s) should exchange information with a lot of other organization and IT system, with whom previously it had no, or only partially had opportunity to come to an agreement, and to create the necessary conditions of IT interoperability. The range of potential cooperation partners spans from the units of the own arm, or own armed force, through the allied, or coalition organizations, to most diverse organizations. At the same time this scale demonstrates the differences in interests, in the closeness of cooperation, in the level of autonomy, and as a consequence in the amount, and characteristics of information exchange relations.

The strategic vision emphasizes the role of information superiority, as a fundamental factor, and the dependence of organizational success on the extensive, and efficient application of information, information processes, information systems, and the services provided by information technology. In the document particularly points out the role, and significance of information (in the first place intelligence) sharing, and creation of situational awareness. [2 (points 14., 18., 31.)] The consequence of this statement is the continuous development in the exploitation of IT systems,

applications, and information handled by them, and in the amount of information exchanged between IT systems of different actors.

Finally one of the most stressed component of the NATO commanders' strategic vision is the emerging network oriented approach that plays a significant role in doctrinal ideas, and its NATO concept, the network enabled capability. [2 (points 29., 32.)] Both on organizational and system level this approach essentially requires an ability to interconnect with other components on a mission-oriented way, to synergically complement each others capabilities, and an ability to efficiently adjust, adapt, and self-reconfigure to a dynamically changing environment.

According to commonly accepted understanding, network centric force is based on the networking of sensors, gathering information; systems, and devices used in mission execution, exploiting information; and command and control systems, and tools supporting organizational level information processing (analysis, evaluation, and decision). This extremely increases information (data) exchange requirements mainly on the level of technical systems, and devices. According to network oriented approach a given IT system should be able to exchange (or acquire) information with (from) existing, and newly appearing systems of a cooperative, neutral, and even adversary actors of infosphere.

As a summary it can be stated, that ideas formulated in the NATO strategic vision describe, outline, and prognose such an information interoperability environment, where:

- conditions of information, and IT interoperability should be ensured for a dynamically extending, and a mission-oriented way changing circle of actors of the international security sphere;
- amount of information handled by the individual actors, and exchanged between them is continuously increasing, its content is dynamically changing;
- more and more increasing part of information appears in IT systems, and is exchanged between them, and in a significant manner extends the amount of connections between IT systems.

All these facts naturally influence the quantity, content, and inner representations of information handled by military IT systems, and the quantity, content, and intermediary representations of information exchanged between IT systems.

Concept and necessity of an interoperability infrastructure

In case of cooperating information systems interconnected with other, heterogeneous systems, transformations between their inner representation and different intermediate representations (languages, message formats) today typically are realized by interface application components related to individual external representations. According to this solution, a new possibility of cooperation, a new intermediate representation requires the development, and implementation of a new interface component.

In a dynamically changing information environment the adaptation based on a continuous development neither sufficiently efficient, nor flexible, and in some cases even can not be accomplished. Even a minor information system upgrade, limited in range and volume, requires a significant amount of time from the formulation of the requirements to the implementation of the new software or hardware version (solution). Moreover an additional time is necessary to do the modifications on all of the working implementations of the given system. What is more, in case of "legacy" systems usually it is not possible to upgrade the system, to extend it with a new interface functionality.

An other disadvantage of interface components connected to individual intermediate representations, is that in case of similar, or identical data elements (e.g. date, time, spatial, quantity characteristics) the transformation between different representations should be repeatedly implemented in different components. This means that knowledge pieces used during transformations are hidden in the individual interface components, so they are can not be re-used.

Network centric approach, widely spreading in our days – that is characterized by the improved accessibility, autonomy, even detachment of different capabilities, and functions earlier

strongly connected to, or inherently built into a "platform" – can be used in case of application components ensuring information interoperability, supporting information exchange among heterogeneous cooperating information systems.

The consequence of the separation of mediator application components from the individual information systems, applications is the implementation of an interoperability infrastructure that is situated between the individual systems, and the communication infrastructure, or constitutes part of an integrated infocommunication infrastructure. So autonomous application components, intended to resolve heterogeneity between different systems (mediators), belong to the group of so-called middleware components.

Information interoperability infrastructure, built on top of a traditional communication system, is a widely available unified system of personnel, devices and services that's purpose is to ensure information exchange based on common understanding, between cooperating IT systems. Basic components of interoperability infrastructure are functional (application) components that implement interoperable transformations.

From the point of view of communication system, interoperability infrastructure is such a value-adder service layer, that works built on top of traditional communication service layers. Basic purpose of communication systems, networks instead of traditional information transmission nowadays is more and more support of integrated functioning of a given organization, community of cooperation as a unified entity. These systems, networks connect particular officials, and organizational units, and they should do this independently of the fact that "what language" the individual actors "speak", what representations they use during information exchange.

Implementation of interoperability infrastructure will probably according to service oriented architecture widely spreading nowadays. In this architecture different components implementing interoperable transformations appear as autonomous service-providers that can be used to build a complex transformation. This kind of implementation, in case of missing interoperability subfunctionality, makes possible to extend the infrastructure with application-side interoperability components, even to provide this functionality to other applications, and later to build into the infrastructure. All this ensures dynamic, application-requirement oriented extension of infrastructural services. At the same time service oriented architecture requires implementation of service-advertisement, -registering, -seeking, and task decomposition/subtask planning functionalities that should also be part of the interoperability infrastructure.

Several components of an interoperability infrastructure are already available in form of independent applications, operation system functions, or functions of system development, and run-time platforms, only conditions of their wider availability should be ensured by placing their services at users' disposal in form of infrastructural services. Components providing interoperability services can be operated in linking nodes of the communication system, in autonomous infrastructural servers, or in application systems at their operation system, or middleware layers.

As it follows from the statements presented earlier, in case of traditional solutions a system has to know a number of different intermediate representation. Whereas in case of an infrastructure-based solution, a given system can use its own, inner representation to send and receive messages, information. Transformation between heterogeneous representations, and transmission of information is the task of the interoperability and communication infrastructure. Because the application components making transformations, are no longer parts of the individual information systems, in addition to the information exchanged, they have to get information (meta-information) about the representations used by the given systems.

Knowledge pieces necessary for understanding (interpretation) of data elements, are called a context. In this sense a context is a system of information (concepts, facts, assumptions, rules) necessary to understand (reproduce) the intended meaning of data elements, data collections belonging to a given information/knowledge representation. According to leading researchers of this topic, a context contains "meta data relating to its meaning, properties (such as source, quality,

and precision), and organization" [3]. Contexts usually can not be fully described, and formalized, because a correct interpretation of information requires the concepts, and rules of "common sense".

In case of traditional implementation of information interoperability, significant amount of context (meta)information needed for interpretation, and understanding is available only in documentations of the given system. Moreover in certain user communities there can be additional interpretation rules that are not recorded in any documentation. On the contrary, an autonomous interoperability layer requires a formalized description of different contexts.

Context descriptions play an interface role, but in contrast with traditional interface components (converters), they are independent of intermediate representations used for information exchange, and they should only have knowledge about the native representation of the system. Another function of context descriptions is supporting inner representation autonomy of the given system. Under appropriate conditions, assuming appropriate capabilities of the interoperability layer, inner representation of a given system, or some parts of it – with simultaneous modification of the context description – can be freely modified.

From the ideas, and suggestions outlined before it seems necessary, and promising to continue researches about a network centric interoperability layer. These should be, and can be coordinated with, among others, the principles, and solutions of the emerging component-based development, and web services architectures. Military sciences should analyse specialities of military application, and should provide subject matter specific knowledge to develop, and implement useable, and useful interoperability solutions, products.

REFERENCES

[1] *The Prague Summit and NATO's Transformation. A Reader's Guide.* – NATO, 2003.

[2] *Strategic Vision: The Military Challenge (by NATO Strategic Commanders).* ACT–ACO, 2004.

[3] SCIORE E., SIEGEL M., ROSENTHAL A.: Using Semantic Values to Facilitate Interoperability among Heterogeneous Information Systems. In. *ACM Transactions on Database Systems.* 1994/2. pp 254-290.

НЕВРОННА МРЕЖА ЗА ОТКРИВАНЕ НА ВОДНИ ТЕЧОВЕ

Сотир Н. Сотиров

NEURON NETWORK FOR DISCOVERY OF WATER LEAKS

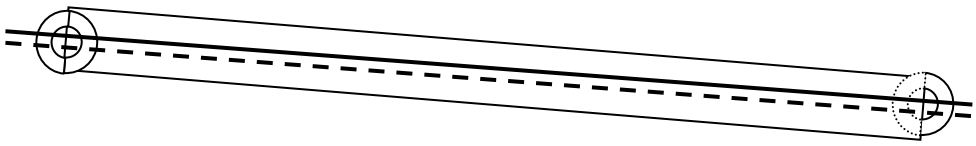
Sotir Sotirov

The article presents a method for discovery and locating water leaks. For the improvement of that method a neuron network has been supplemented, which allows the unveiling of cracks in the network when there are flood and condensation

KEY WORDS: neuron network, diagnostics of water-supply network

Въведение:

Във всяка жилищна и административна сграда има водопроводна инсталация. Течовете предизвикани от спукването на тръби не са рядкост. За откриването на повредите са използвани множество способи [4-7].



Фиг.1

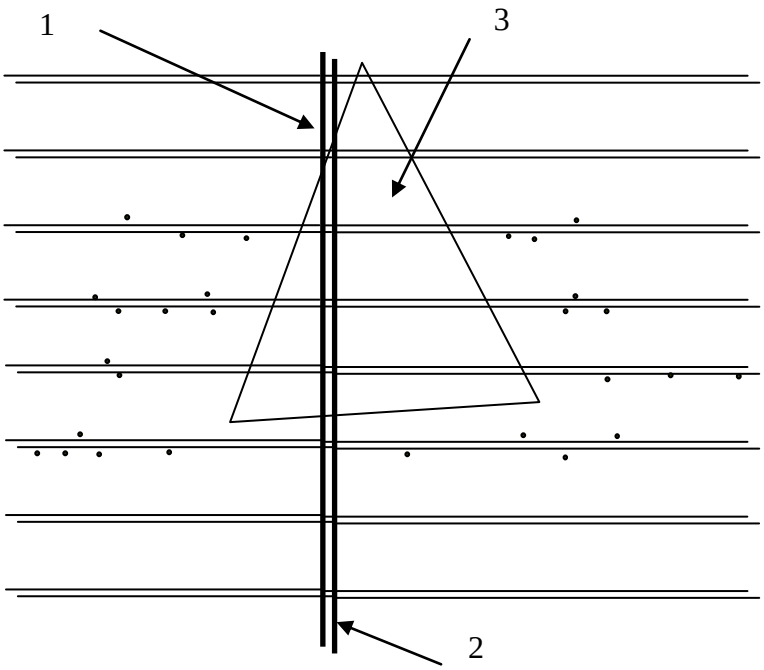
Една от технологиите използва влакно съдържащо два проводника [4]. Единият от проводниците е меден, а другият е със съпротивление. При проникване на капка от спукване на канализацията във влакното то дава контакт между двата проводника.

Съпротивлението R получено при този процес е пропорционално на разстоянието от началото на влакното. Колкото по-близо е теча до началото на влакното толкова по-малко е съпротивлението R.

За определяне на местоположението на теча е необходимо да се измери съпротивлението между двата проводника.

В класическия случай за откриване на мястото на течът се използва мост на Уинстон [4]. За едно от съпротивленията се използва цифрово управляем резистор. С помощта на микропроцесорна система Уинстоновият мост се уравновесява. Така цифровото показание на входа на управляемият резистор дава местоположението на контакт между двата проводника.

Много често при спукване на тръба са получава миниатюрен разлив, при което се получават множество контакти между проводниците. Тогава при използването на горния метод се отчита контактът между двата проводника който е най-близо до началото на влакното (от страната на микропроцесорната система). Това предизвиква погрешно отчитане на местоположението на спукването на тръбата, породено от разливането на течността (фиг.2).



Фиг. 2

Схема на разположение на водопровода

1. Водопровод за топла вода
2. Водопровод за студена вода
3. Област на отчитане на разлива

Нека за постигането на по-добра точност на отчитането на местоположението на спукването се постави второ влакно. За да може да се отчете с по-голяма точност местоположението на пробива второто влакно е с посока обратна на първото.

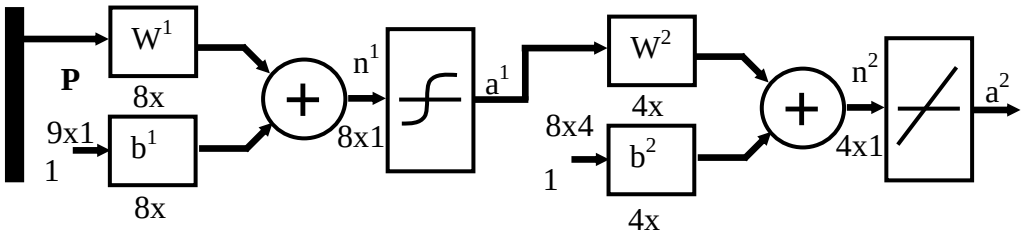
За по-точното откриване на местоположението е целесъобразно да се използва невронна мрежа. С нея при комплексните разливи и конденз може да се открие първоизточника на водния теч.

Нека всяка двойка влакна след преобразуването в цифров еквивалент, подава на входовете на невронна мрежа двойка стойности, показващи наличието на теч и разположението му спрямо началото на влакното. Нека със стойности 0 да се отчита липса на контакт между двата проводника, а със стойности от 0.001 до 1 местоположението на контакта.

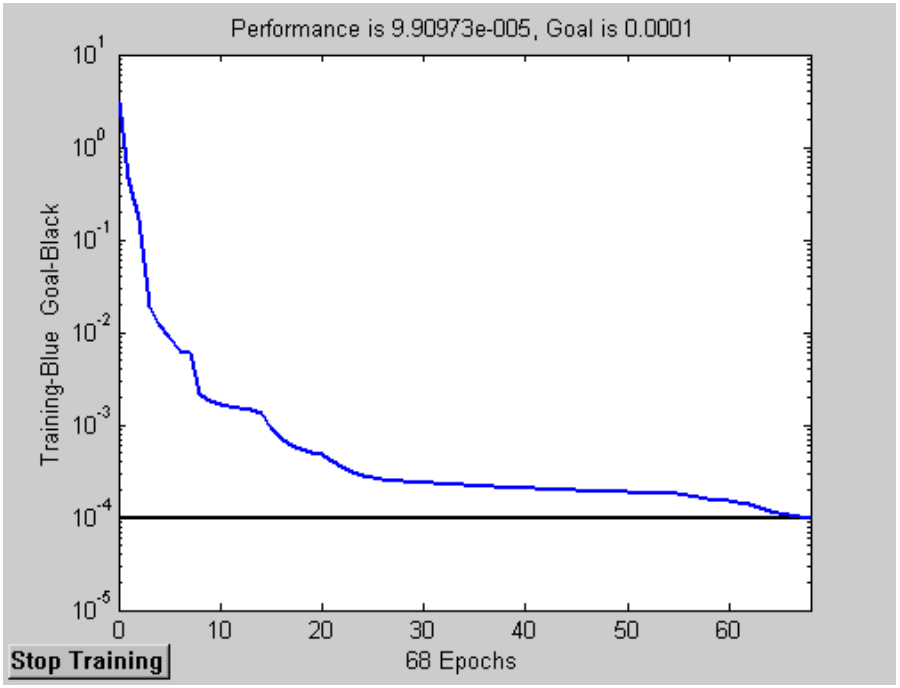
Реализация

За реализацията е използвана трислойна невронна мрежа със структура 9:8:4 (фиг.3). На осемте й входа (за четири двойки влакна) се подават двойките стойности, а на деветия вход са подава информация за влажността на въздуха. Вторият й слой се състои от 8 неврона, като за трансферна функция се използва логическа функция (logsig). Изходният слой е с линейна трансферна функция.

Обучението на мрежата е извършено в среда на MATLAB по алгоритъма Левенбенг-Маркуард [8, 9] (вариант на алгоритъма ”обратно разпространение” [10]) със зададена средна квадратична грешка 1.10^{-4}. При обучението за изход се задават стойности, показващи местоположението на спукването или липсата на такова. Графиката на изменението на грешката при обучението е показана на фиг.4.



Фиг.3



Фиг.4

Стойностите на тегловните коефициенти W^1, W^2 и отместванията b^1 и b^2 са показани по долу.

W_1 =

-1.6394	-14.844	3.6327	-12.7501	-7.4498	2.1070	0.7962	3.5808	-10.5769
4.3503	2.6482	2.3925	6.6088	-3.2302	2.2588	1.9933	1.9807	-0.9816
-1.2374	19.218	-5.4135	-2.4678	5.1085	-0.9330	-0.8024	0.9888	9.1874
-0.7590	-1.5444	1.8993	-1.6816	0.2257	15.0742	-3.3277	1.2620	-0.1626
5.5810	0.1671	2.2065	-8.8721	-7.5484	-10.2836	-9.0895	-8.1527	-7.4097
-1.2036	1.7561	-2.5488	-2.3877	2.6970	-4.7656	-3.5880	4.4337	2.0603
0.3269	-2.1422	3.0867	0.7367	-1.6566	5.8328	4.2244	-1.5496	-1.5233
-1.1518	5.3128	-9.3636	-0.0289	1.4195	3.7676	-9.3498	-2.2266	2.2522

b_1^t = 8.6780 -8.2583 0.2571 1.6485 11.7989 -0.3797 0.2464 6.4960

W2 =

0.2862	2.8119	1.6419	1.5698	-0.6196	-2.1424	-1.5676	-0.2168
0.5618	-1.9645	-1.1287	-1.1719	0.3824	-1.3893	-2.4403	0.1616

-0.3903 1.3467 -0.3763 2.3511 3.7543 -0.4873 -0.4582 0.9004
0.2606 -0.0529 -0.2487 -8.4699 -0.1012 7.6199 9.9306 0.6848

b_2^t = -0.1062 2.9603 -5.4176 -2.2599

Невронната мрежа се обучава за 68 итерации (фиг.4).

Тестване

При тестването на входа на невронната мрежа се подават 10 тест вектора комбинирани така, че да показват различните характеристики на системата.

Първият тест вектора е за нормално (незадействано) състояние на невронната мрежа.

Следващите три вектора са така зададени като случайни контакти само в едно от влакната имитирайки конденз предизвикан от висока влажност на въздуха.

Следващите три вектора са за тест при откриване на реални течове.

Последните три са вектори, с които мрежата е била обучавана.

№	Входен тестов вектор (P)									Изходи на невронната мрежа			
1	0	0	0	0	0	0	0	0	0	0.1922	0.0159	0.1892	0.1470
2	0	0.8000	0	0	0	0	0	0	0	-0.1646	0.0089	0.0508	0.0293
3	0	0	0	0	0	0.1000	0	0	0	0.1778	0.1878	0.0578	0.1259
4	0	0	0.2000	0	0	0	0	0	0	0.0720	0.0159	0.1328	0.1123
5	1.0000	0.8000	0.2000	0	0	0	0	0	0	0.8010	0.0011	0.0005	0.0021
6	1.0000	0	0	0.4000	0.3000	0	0	0	0	-0.0986	1.0767	0.2222	0.0248
7	1.0000	0	0	0	0	0	0	0.1900	0.7900	0.0133	0.0093	0.2530	0.0636
8	0	0	0	0.2000	0.2000	0	0	0	0	-0.0021	0.5059	0.0286	0.0114
9	1.0000	0	0	0.4000	0.6000	0	0	0	0	-0.0010	0.3990	0.0029	0.0010
10	1.0000	0	0	0	0	0	0	0.7000	0.3000	-0.0009	0.0065	0.0001	0.6953

Изводи

На базата на изследванията може да се направят следните изводи:

- Невронната мрежа разширява стандартната електронна система за откриване на течове;
- Получената система е в състояние да открива местоположението на спукване на база на две противоположно разположени влакна и по този начин се намира широчината на разлива;
- При подаване на нулеви стойности на входа на мрежата всички изходи са с приблизително еднакви стойности (ред едно от таблицата);
- При отчитане на отделни произволни стойности само в едната от противоположните влакна не се отчита спукване и съответно разлив и се дължат на конденз (редове 2-4);
- При подаване на стойности показващи реални спуквания и разливи (редове 5-7) на изхода на мрежата се получават стойности показващи местоположението на разлива и на кое рамо е - / ред 7 последния клон е с отчети 0.19 и 0.79, т.е разлива е на 0.2 от едната страна и на 0.8 от другата страна/;

Заключение

Предложената невронна мрежа увеличава възможностите за откриване на течове в големите обществени сгради използвайки стандартни техники за измерване. Тя може да се използва и в случаите, при които влажността на въздуха е с висок процент и се създават отделни капчици разположени произволно по мрежата. Още по-сложният вариант е комбинацията от двете (конденз и разлив) във водопроводната система на голяма обществена сграда не би могла да се открия и съответно да се локализира по класическия вариан.

ЛИТЕРАТУРА

1. **Haykin**, S. “Neural Networks: A Comprehensive Foundation”, Prentice Hall, N.J., 1999
2. **Bishop** C. M., Neural networks for pattern recognition, Oxford university press, ISBN 0 19 853864 2 , 2000
3. **Hagan**, M.T. H.B.Demuth, M.Beale, “Neural Network Design”, PWS Publishing Company, Boston, 1996
4. **Shark** Marine Technologies Inc., http://www.sharkmarine.com/locating.htm
5. **Metrotech** LTD, www.aegis.net.au
6. **Marian Trela**, Dariusz Butrymowicz, Jerzy Głuch, Andrzej Gardzilewicz, Eugeniusz Ihnatowicz, MONITORING OF AIR CONTENT IN A MIXTURE REMOVED FROM CONDENSERS IN APPLICATION TO STEAM TURBINE DIAGNOSTICS, Proceedings of 2000 International Joint Power Generation Conference Miami Beach, Florida, July 23-26, 2000 IJPGC2000-15002
7. **Logis-Tech Associates**, A-1 Technical Training, http://www.logis-tech.co.uk/
8. **Levenberg**, K. "A Method for the Solution of Certain Problems in Least Squares." Quart. Appl. Math. 2, 164-168, 1944.
9. **Marquardt**, D. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters." SIAM J. Appl. Math. 11, 431-441, 1963.
10. **Rumelhart**, D. E., G. E. Hinton and R.J. Williams, “Learning representation by back-propagation errors”, Nature, vol.323, pp 533-536, 1986.

ОСНОВНИ МЕТОДИ ЗА ЗАРАЗЯВАНЕ НА КОМПЮТЪРНИТЕ СИСТЕМИ С ВИРУСИ

АНГЕЛ В. КОЛЕВ, НИНА СИНЯГИНА

MAIN METHODS FOR INFECTING OF THE COMPUTER SYSTEMS WITH VIRUSES

ANGEL V. KOLEV, NINA SINIAGINA

The present material show the difference ways that are using for infecting of the computer systems using in the global network Internet. The better analyzing of the methods using for infecting of the systems is the main step for making of new methods of neutralizing of the viruses attacks.

KEY WORDS: virus, infection, method, Internet, computer.

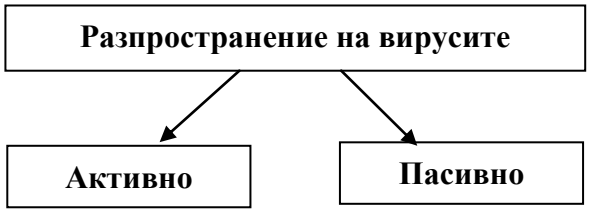
1. ВЪВЕДЕНИЕ

С развитието на Интернет и свързването на голяма част от компютърните системи в една глобална мрежа освен предимствата, които се създават с бързия достъп на информацията от цял свят, се създават предпоставки за бързо разпространение на вирусите. Компютърните вируси, чрез глобалната мрежа Интернет, се разпространяват за няколко часа до всяка точка на света, използвайки различни методи на заразяване, които авторите са разгледали в представения материал.

2. ИЗЛОЖЕНИЕ

С развитието на глобалната мрежа Интернет и развитието на все по-бързите комуникации, използващи Интернет пространството разпространението на вирусите се превръща в основен елемент в областта на сигурността на комуникациите. Вирусите за няколко часа успяват да заразят хиляди компютърни системи по цял свят и да блокират дейността на редица организации и фирми. Тяхното разпространение е дейност както на хора, занимаващи се с кражба на информация чрез Интернет, но също така и на хакери и чисто експериментални действия от страна на програмисти, занимаващи се със сигурността на компютърните мрежи.

Разпространението на вирусите може да се раздели на две: едното е активното разпространение на вирусите, когато целенасочено вируса успява да проникне в компютърната система [Фиг. 1], а другото е пасивно разпространение. Пример за първият случай е вируса “Bluster”, който активно се разпространи в края на 2003-та година и зарази всички компютърни системи, работещи в Интернет мрежата с операционната система на Microsoft – Windows XP.



Фиг. 1 – Разпространение на вирусите

Пасивното разпространение е втория начин на разпространение на вирусите. То използва разсеяността на потребителите за да успее да се разпространи. Използват се

различни методи за заблуда на потребителя, при което самия потребител да изтегли вирусния файл на хард диска на системата и да го стартира. Това пасивно разпространение на вирусите се осъществява най-вече от безизвестни Интернет сайтове, разполагащи с интересна, но и съмнителна информация, като се предоставя на потребителя полезен за него софтуер. След изтеглянето на този файл съществуват две основни възможности за заразяване. Първият начин е да се зарази компютърната система директно след инсталацията. В този случай вируса се съдържа в изтегленият файл. Този подход на заразяване е стар метод, който в днешно време по-рядко се използва. Основният недостатък на този метод на заразяване е, че вирусът много бързо и лесно се открива от антивирусните програми. Вторият метод на инфектиране на компютърните системи с вируси от вече изтеглени файлове е метода, използващ комуникацията между компютърната система и Интернет пространството, с което комуникира ново инсталираната програма. При този метод антивирусната програма като сканира файла не открива вируси в него и програмата след позволението, а в някои случаи и без позволение от страна на потребителя, започва своята комуникация в Интернет. Такива програми могат да бъдат програми, свързани с най-различни области, като например да предоставят на потребителя информация за музика, игри, новини, а също така достъп до сайтове с различни забранени съдържания и други. При самото осъществяване на тази комуникация наред с получената информация на потребителя от страна на програмата се предоставя и скрита за него информация – това е комуникационния метод на инфектиране [Фиг. 2]. Тази скрита информация представлява вирусния файл, който достига до компютъра на потребителя, самоинсталира се и започва да функционира и да работи за това, за което е бил създаден. При този метод единственото противодействие е наличието на разидентна антивирусна програма, за да бъде открит този файл и да се пристъпи към изтриване на вируса. В някой случай след откриването на вируса програмата остава скрита и не е възможно да се определи какъв е пътя на инфектиране. В този случай блокирането и спирането на вирусите става единствено чрез откриването на първоизточника за разпространение на вируса. Един от основните начини за откриване на програмата, която е създала условия за навлизане на вируса в компютърната система е не използването на антивирусна програма, а самия оператор. Той трябва да определи коя от програмите, които са стартирани е направила съмнителна връзка в Интернет и от къде е дошъл вируса. В този случай вируса може да се помести не в директорията на тази програма, а в системните директории на операционната система. Подпомагането на дейността на потребителя, който в този случай трябва да изпълнява дейността на системен оператор, трябва да открие вируса и да спре понататъчното разпространение е наличието на защитна стена. Без използването на защитна стена откриването и спирането на разпространение на вирусите става значително по-сложно за потребителите.

Фиг 2. Методи на инфектиране на компютърните системи

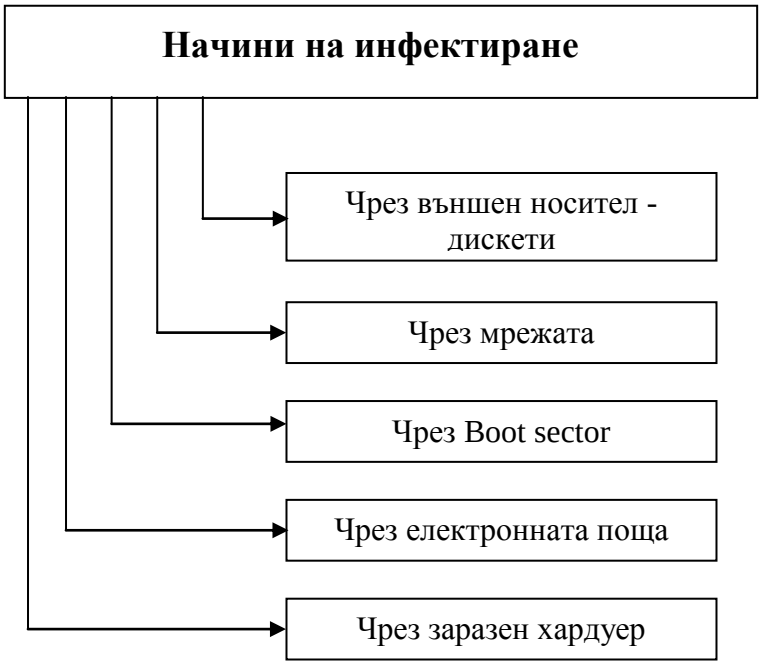
Тези три метода на инфектиране на компютърните системи са основните, при които се осъществява проникване на вирусите в компютърната система. Противодействието на тези методи на инфектиране са няколко. Най-важното условие е наличието на резидентна

антивирусна програма, която да е активна дори и по време на инсталацията. Като второ необходимо условие е наличието на защитна стена, която да блокира достъпа на тези програми, които се опиват да комуникират с Интернет пространството. Тези две условия не гарантират сигурността на една компютърна система, но са условия да се повиши в значителна степен защитата на системата от атаките на вирусите.

Инфектирането на една система освен чрез използването на Интернет мрежата се осъществява и чрез следните начини [Фиг. 3]. Тези методи на инфектиране са основно пет. Един от начините на инфектиране се осъществява чрез външен носител, като дискети и дискови устройства. Този метод на инфектиране е стар метод, който се използваше през последните години на миналия век. Следващият метод на инфектиране е чрез мрежата. Това е сравнително нов метод на инфектиране, който използва пропуските в сигурността. Неговата ефективност е много висока и е метода, който най-често ще се използва за сигурно заразяване на компютърните системи. Друг метод на заразяване на компютърните системи е чрез Boot Sector. Този метод се използва сравнително рядко, но тези вируси са едни от най-опасните вируси при, които изтриването им се осъществява само чрез форматиране.

Метода, който се използва най-много през последните няколко години е заразяване чрез електронната поща. Този метод позволява, чрез най-използваната услуга в Интернет да се осъществява бързо инфектиране на компютърни системи от цял свят за сравнително малко време. При метода на инфектиране чрез електронна поща, който използва както заблуда на потребителите, така и в някои случаи заразяване на компютърната система чрез директното попадане на вируса на хард диска на компютъра.

Последния начин на инфектиране на една компютърна система е чрез използването на заразен хардуер. При слагане на втори хард диск или при инсталирането на допълнителна памет към една компютърна система също може да доведе до инфектиране на системата.



Фиг 3. Начини на инфектиране на компютърната система

Заразяването на компютърните системи непрекъснато се променя, създават се нови методи и модели на атака. Освен експериментални и хакерски атаки, те биват и целенасочени атаки, дължащи се на отделни фирми и организации, които се опитват да придобият информация за конкуренти или извършването на

финансови престъпления чрез интернет мрежата. Всичко това е подчинено на начините на заразяване на компютърните системи. През последните години разпространението на вирусите е активно, като се използват комуникационния метод и метода чрез самоинсталиране. Начините на инфектиране са предимно два - чрез мрежата и чрез електронната поща.

3. ЗАКЛЮЧЕНИЕ

Вирусите са основният проблем на всички потребители, които работят с глобалната мрежа. В настоящият материал са представени методите на разпространение на вирусите и методите на инфектиране. С извеждането на проблемните точки и начините, по който се осъществява заразяването на компютърните системи се определят най-уязвимите начини, които се използват за атака от страна на вирусите. Борбата между антивирусните методи и методите на атака от страна на вирусите ще продължи и за в бъдеще да бъде основен проблем за сигурността на компютърните системи, намиращи се в глобалната мрежа Интернет.

ЛИТЕРАТУРА

1. **Петров** Р., "Защита на информацията в компютрите и мрежите", София.
2. **Tanenbaum** S. Andrew , Modern Operating Systems, 2nd edition, New Jersey, Prentice Hall, 2001.
3. http://www.cnet.com
4. http://www.pandasoftware.com

УСКОРЕНИ АЛГОРИТМИ ЗА ГЕНЕРИРАНЕ И РАБОТА В ПРОСТО ПОЛЕ НА ГАЛОА *GF*(*P*)

ЖАНЕТА Н. ТАШЕВА

FAST ALGORITHMS FOR GENERATION AND CALCULATION IN GALOIS FIELDS *GF*(*P*)

ZHANETA N. TASHEVA

In the paper first the briefly review of the basic mathematic background of Galois Field GF(p) is given. Then the fast algorithms for generation and calculation in Galois fields are proposed. Finally the software implementation of proposed algorithms is discussed. The suggested fast algorithms and program model allow generation of Galois Field with p up to 2^{32} and polynomial calculations in generated GF(p).

KEY WORDS: Galois Field, GF(p), Fast Algorithm, Galois Field Generation

I. Увод

Предавателните канали, особено безжичните, не са идеални поради наличието на много смущения (шумове, интерференции, ехо ефекти), които се наслагват с полезния сигнал и водят до поява на грешки. За да се осигури добро качество на приеманото изображение, е необходимо да се осигури много ниско ниво на грешките в канала. Например в системата *DVB*[8], е необходимо да се постигне *BER*[9] от порядъка $10^{-10} - 10^{-12}$, при скорост на предаване на данните 30 *Mb/s*. Канал с толкова ниско ниво на грешки се нарича *квази-безгрешен* канал. За да се постигне това ниско ниво на грешки е необходимо да се вземат мерки, чрез които в приемната страна да се осигури детектиране и коригиране на колкото е възможно повече грешки. Това може да се постигне чрез шумоустойчиво кодиране, като се въведе пресметнат излишък информация. Едни от най-често използваните при прогресивната корекция на грешки шумоустойчиви кодове са кодовете на Рид-Соломон (*RS*[10]) и Боуз-Чоудхури-Хоквингем (*BCH*[11]). Те позволяват корекция на въведените от предавателния канал пакети от грешки.

Едно от съвременните направления за работа в тази област използва идеята за *адаптивна промяна на използвания шумоустойчив код* в зависимост от получените грешки в канала. Това води до изискването за използване на повече от един код, коригиращ пакети грешки и бърза пренастройка на кодера на канала към новоизбрания код. Решаването на тази задача може да се осъществи чрез програмно изпълнение на задачите по кодиране и декодиране на използваните кодове. Допълнително е необходимо да се реши задачата за генериране на нов шумоустойчив код, коригиращ друг брой грешки в зависимост от текущите характеристики на канала. Понеже *RS* и *BCH* са линейни кодове, т.е. *прости векторни пространства,* които съществуват единствено, ако размерът на азбуката q е просто p или степен на просто число p^m, за генериране на кодовете първо е необходимо да се синтезират *ускорени алгоритми* за генериране и работа в поле на Галоа $GF(p)$[12] и негово разширение $GF(p^m)$.

Актуалността на проблема за *информационната защита* на предавателните канали се засилва през настоящото хилядолетие от необходимостта за постигане на успех в информационната война, която ще се води с изключителна острота в кибернетичното пространство при всеки кризисен или въоръжен конфликт. Дори и в мирно време, е налице непрекъсната надпревара за създаване на нови методи и средства за дезинформация, за дешифриране на скрита информация, за дестабилизиране на

[8] ***DVB (Digital Video Broadcasting)*** – Стандартизирана система за цифрови видео разпръскване
[9] ***BER (Bit Error Rate)*** – Ниво на битови грешки
[10] ***RS (Reed - Solomon)*** - Цикличен шумоустойчив код на Рид–Соломон
[11] ***BCH (Bose-Chaudhuri-Hocquenghem)*** – Цикличен шумоустойчив код на Боуз–Чоудхури– Хоквингем
[12] ***GF(p) (Galois Field)*** – Поле на Галоа с характеристика p

информационните системи на потенциалните противници. Това налага необходимостта от създаване на *ускорени алгоритми* за генериране и работа в поле на Галоа $GF(p)$, които са математическата основа за изграждане на *LFSR*[13] регистрите в *PRNG*[14].

Това определя все по–нарастващото приложение на аритметиката в полетата на Галоа $GF(p)$ и $GF(p^m)$ при съвременните приложения за шумоустойчиво кодиране [7] и в криптографията за генериране на псевдослучайни последователности с голям период на повторение [6, 8].

В настоящата статия първо е направен обзор на математическата основа на простото поле на Галоа $GF(p)$. След това са предложени ускорени алгоритми за генериране и работа в $GF(p)$. Накрая е представена софтуерната реализация на синтезираните ускорени алгоритми.

II. Математическа основа на просто поле на Галоа $GF(p)$

В тази част на доклада накратко е обобщена структурата на простите полета на Галоа $GF(p)$. Изложени са някои основни определения и теореми от математическата теория на крайните полета, използвани в представения в следващата част на доклада алгоритъм. По-подробно описание на теорията на крайните полета може да се намери във всяка книга, разглеждаща проблемите на съвременната алгебра [1, 2, 3] и алгебричната теория на кодирането [4, 5, 7].

Определение 1: Крайните полета съществуват само в случаите, когато броят на елементите им е просто число или степен на простото число. В първия случай полето се нарича *просто*, а във втория – *разширение*, съответстващо на простото поле [1, 2, 5].

Определение 2: Простото поле на Галоа $GF(p)$, където p е просто число, е крайно множество от p на брой различни елемента – $0, 1, \ldots, p-1$ [1, 2, 5].

Определение 3: Порядък (характеристика) на полето се нарича броя на неговите елементи. Следователно порядъкът на поле на Галоа $GF(p)$ е p [1, 2, 5].

Основните свойства на полето на Галоа $GF(p)$ могат да се обобщят по следния начин:

1. Определени са две основни операции: *събиране* (+) и *умножение* (.):

Ако a и b са произволни елементи от $GF(p)$, то тяхната сума c и произведението им d се дефинират по следния начин.

(1) $$c \equiv (a + b) \bmod p$$

(2) $$d \equiv (a \,.\, b) \bmod p$$

2. Резултатът от сбора c и умножението d на два произволни елемента от полето [$a \in GF(p)$; $b \in GF(p)$] е трети елемент, който също принадлежи на това поле [$c \in GF(p)$; $d \in GF(p)$].
3. Полето на Галоа $GF(p)$ съдържа нула 0 (адитивна единица) и единица 1 (мултипликативна единица), такива че за всеки произволен елемент a, са в сила равенствата:
$$a + 0 = a$$
$$a \,.\, 1 = a.$$
4. За всеки елемент a съществува обратен елемент при събиране ($-a$) и обратен елемент при умножение a^{-1}, като са изпълнени равенствата:
$$a + (-a) = 0$$
$$a \,.\, a^{-1} = 1.$$

Съществуването на тези обратни елементи, позволява да се изпълняват действията изваждане и делене в крайните полета.

5. В полето се изпълняват следните закони:

(3)
1. комутативен закон $a + b = b + a$;
2. асоциативен закон $(a + b) + c = a + (b + c)$, $(ab)c = a(bc)$;
3. дистрибутивен закон $(a + b)\, c = a\, c + b\, c$, с $(a + b) = ca + cb$.

Пример за резултатите от операциите събиране и умножение в полето $GF(11)$ са дадени съответно в таблици 1 и 2. За изпълнение на операциите изваждане и деление, е необходимо първо да

[13] ***LFSR (Linear Feedback Shift Register)*** – Линеен регистър с обратни връзки

[14] ***PRNG (Pseudo Random Number Generator)*** – Псевдо–случаен генератор

се намери съответния обратен елемент от таблицата и след това да се извърши съответното събиране или умножение. Например в *GF*(11):

$$5 - 7 = 5 + (-7) = 5 + 4 = 9.$$

Аналогично:

$$5 / 7 = 5 . (7^{-1}) = 5 . 8 = 7$$

Таблица 1. Събиране в *GF*(11)

+	0	1	2	3	4	5	6	7	8	9	10
0	0	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10	0
2	2	3	4	5	6	7	8	9	10	0	1
3	3	4	5	6	7	8	9	10	0	1	2
4	4	5	6	7	8	9	10	0	1	2	3
5	5	6	7	8	9	10	0	1	2	3	4
6	6	7	8	9	10	0	1	2	3	4	5
7	7	8	9	10	0	1	2	3	4	5	6
8	8	9	10	0	1	2	3	4	5	6	7
9	9	10	0	1	2	3	4	5	6	7	8
10	10	0	1	2	3	4	5	6	7	8	9

Таблица 2. Умножение в *GF*(11)

*	0	1	2	3	4	5	6	7	8	9	10
0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9	10
2	0	2	4	6	8	10	1	3	5	7	9
3	0	3	6	9	1	4	7	10	2	5	8
4	0	4	8	1	5	9	2	6	10	3	7
5	0	5	10	4	9	3	8	2	7	1	6
6	0	6	1	7	2	8	3	9	4	10	5
7	0	7	3	10	6	2	9	5	1	8	4
8	0	8	5	2	10	7	4	1	9	6	3
9	0	9	7	5	3	1	10	8	6	4	2
10	0	10	9	8	7	6	5	4	3	2	1

Определение 4: Порядък на елемента *a* от полето се нарича броя на различните степени на този елемент [1, 2, 5].

Всички крайни полета, в това число и полетата на Галоа *GF*(*p*), съдържат елемент, наречен *генератор на полето* или *примитивен елемент*. В сила е следната теорема [2, 4, 7]:

Теорема 1: Крайното поле от ред *p* съдържа примитивен елемент α от порядък $p - 1$, степените на който обхождат всички елементи на полето:

(4) $$\alpha^{p-1} \equiv 1 \bmod p$$

Например, както се вижда от таблица 2, в полето *GF*(11) $\alpha = 2$ е примитивен елемент, защото:

$2^0 = 1$; $2^1 = 2$; $2^2 = 4$; $2^3 = 8$; $2^4 = 5$; $2^5 = 10$; $2^6 = 9$; $2^7 = 7$; $2^8 = 3$; $2^9 = 6$; $2^{10} = 1$;

Полетата на Галоа от порядък $p > 3$ имат повече от един примитивен елемент. От математическа гледна точка намирането на един такъв елемент α е достатъчно за извършване на всички действия в полето. Но за целите на алгебричното кодиране и криптографията е от значение намирането на всички примитивни елементи на полето, защото те позволяват да се получат други пермутации на елементите на полето, представени като степени на примитивния елемент α.

Например, използването на примитивния елемент $\alpha = 6$ позволява да се получи следната пермутация на елементите от полето *GF*(11):

$6^0 = 1$; $6^1 = 6$; $6^2 = 3$; $6^3 = 7$; $6^4 = 9$; $6^5=10$; $6^6 = 5$; $6^7 = 8$; $6^8 = 4$; $6^9 = 2$; $6^{10}= 1$;

Традиционният алгоритъм за намиране на всички примитивни елементи в поле на Галоа *GF*(*p*) включва последователно обхождане на всички елементи $a \geq 2$ от полето и проверка дали те са от порядък $p - 1$ (теорема 1). Следователно за извършването му са необходими $(p-1)(p-2)$ на брой операции умножение в *GF*(*p*). При големи стойности на *p* броят на необходимите операции за този алгоритъм нараства многократно, като се има предвид, че операцията умножение (2) в *GF*(*p*) се състои от две операции: умножение и привеждане по модул *p*.

III. Синтез на ускорени алгоритми за генериране и работа в просто поле на Галоа *GF*(*p*)

Ускореният алгоритъмът за генериране на простото поле на Галоа GF(p) *включва четири основни стъпки:*

Стъпка 1: *Проверка дали характеристиката* **p** *е просто число. Ако не, изход.*

Стъпка 2: Намиране на първия примитивен елемент α_1: $\alpha_1^{p-1} \equiv 1 \bmod p$.

Стъпка 3: *Представяне на всеки елемент* α_i *на полето* **GF(p),** *като степен на примитивния елемент* α_1: $\alpha_i = \alpha_1^{\,i} \bmod p$, $i = 0 \ldots p-1$.

Стъпка 4: *Намиране на останалите примитивни елементи:* $\alpha_1^j \bmod p$, *където* j *е взаимно просто с* $p - 1$.

При реализацията на стъпка 1, *за изпитване делимостта на дадено положително число (характеристиката* p*) е използвана теорията на числата, при което изчисленията за непосредствена проверка са ограничени до делители, не надминаващи* $\sqrt{p}$.

При реализацията на стъпка 2, *поради простотата на действията в простите полета на Галоа, намирането на първия примитивен елемент в полето се извършва чрез непосредствена проверка. Опитно е установено, че обикновено* $\alpha_1 = 2$ *или 3, от където започва проверката.*

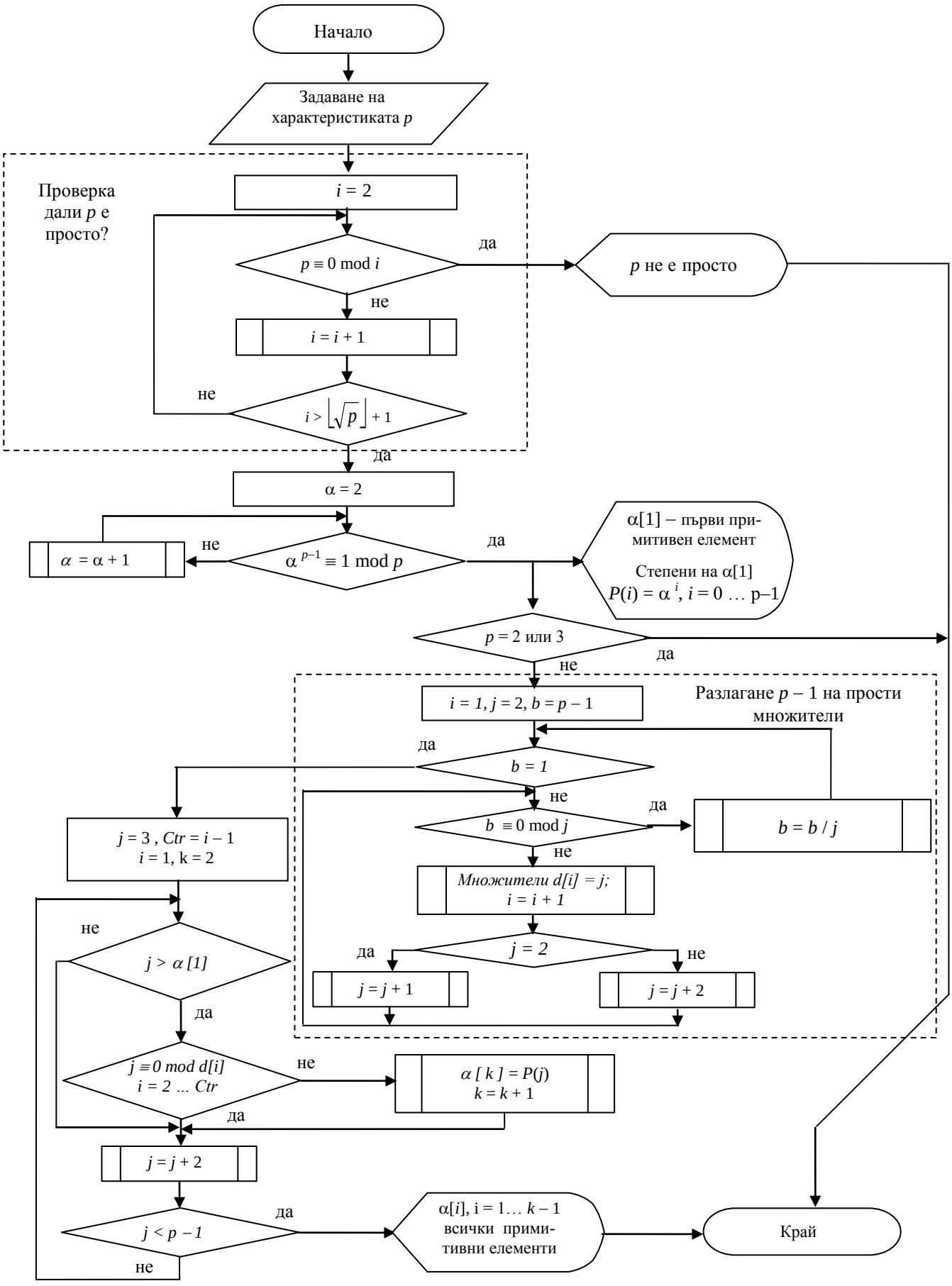
Блоковата схема на предложения ускорен алгоритъм е представена на фиг. 1.

Пример: **Построяване на *GF*(11)**

1. **Първи примитивен елемент $\alpha_1 = 2$, $2^{10} \equiv 1 \bmod 11$**
2. **Представяне на всеки елемент α_i на полето като степени на 2:**

$2^0 = 1$; $2^1 = 2$; $2^2 = 4$; $2^3 = 8$; $2^4 = 5$; $2^5=10$; $2^6 = 9$; $2^7 = 7$; $2^8 = 3$; $2^9 = 6$; $2^{10}= 1$;

3. **Намиране на останалите примитивни елементи:**
   - **разлагане $p - 1 = 10$ на прости множители: $10 = 2.5$;**
   - **$2^j \bmod 11$, $j = 3$, взаимно-просто с 10 $\Rightarrow \alpha_2 = 2^3 \bmod 11 = 8$;**



Фиг. 1. Блокова схема на алгоритъма за генериране на просто поле на Галоа GF(p)

- **2^j mod 11, j = 7, взаимно-просто с 10 $\Rightarrow \alpha_2 = 2^7$ mod 11 = 7;**
- **2^j mod 11, j = 9, взаимно-просто с 10 $\Rightarrow \alpha_2 = 2^9$ mod 11 = 6;**

Оценка на бързодействието: **В оценката не е включено времето за изпълнение на стъпка 1, защото тя се изпълнява както при традиционния, така и при ускорения алгоритъм. Ако построяването на простото поле на Галоа се извърши по пътя на общото търсене ще са необходими общо $(p-1)(p-2)$ умножения в $GF(p)$. При предложения алгоритъм итерациите се намаляват на $\leq p-3$, за стъпки 3 и 4. Опитно е определено, че първият примитивен елемент е по-малък или равен на 10, така че бързодействието на стъпка 2, е винаги $\leq 10(p-1)$. В таблица 3. са представени примери за броят операции при генериране на просто поле на Галоа по традиционния метод на непосредствена проверка и ускорения алгоритъм.**

Тази оценка на бързодействието, дава основание да се твърди, че предложеният алгоритъм многократно подобрява бързодействието като с увеличаване на p бързодействието се увеличава експоненциално, за сметка на намаления брой проверки с цел елиминиране на елементите, които не могат да бъдат примитивни.

Таблица 3. Необходим брой операции за генериране на GF(p)

характеристики	метод	брой проверки	умножения в $GF(p)$	деления
$p = 5, \alpha_1 = 2$	**традиционен**		**4.3 = 12**	
	ускорен	**2**	**4 + 3 = 7**	**2**
$p = 7, \alpha_1 = 3$	**традиционен**		**6.5 = 30**	
	ускорен	**3**	**2.6 + 5 = 17**	**3**
$p = 11, \alpha_1 = 2$	**традиционен**		**10.9 = 90**	
	ускорен	**6**	**10 + 9 = 19**	**4**
$p = 13, \alpha_1 = 2$	**традиционен**		**12.11 = 132**	
	ускорен	**6**	**12 + 11 = 23**	**3**
$p = 17, \alpha_1 = 3$	**традиционен**		**16.15 = 240**	
	ускорен	**8**	**2.16 + 15 = 47**	**4**
$p = 19, \alpha_1 = 2$	**традиционен**		**18.17 = 306**	
	ускорен	**7**	**18 + 17 = 35**	**4**
$p = 23, \alpha_1 = 5$	**традиционен**		**22.21 = 462**	
	ускорен	**11**	**3.22 + 21 = 87**	**6**

Тъй като операциите умножение в GF(p) *включват умножение, последвано от привеждане на резултата по модул на характеристиката* p*, то за ускоряване работата в полето след неговото генериране могат да се пресметнат всички възможни произведения и да се запишат в таблица. След това при работа, резултатът се получава директно чрез индексиране на таблицата.*

IV. Програмна реализация на предложените алгоритми

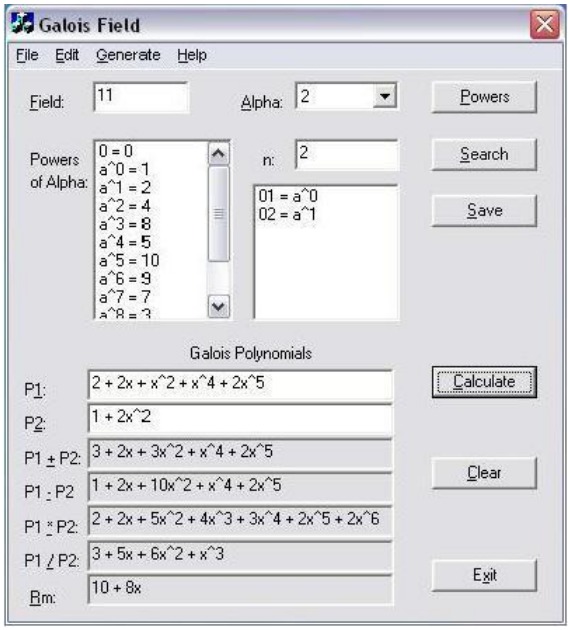
Предложените ускорени алгоритми са залегнали в основата на програмна реализация на свойствата на простите полета на Галоа и работата в тях на езика Visual C++ 6.0*. При реализацията са използвани принципите на обектно ориентираното програмиране, като е създаден* ***основен клас*** **CGaloisField** *(листинг 1). Той има три* основни свойства*: характеристика* **p** *(Cardinality), примитивен елемент* α *(Alpha) и режим на изобразяване на екрана DisplayMode.*

Листинг 1: *Основен клас* CGaloisField

```
class CGaloisField : public CObject
{
  public:
        void SetAlpha(int newalpha);        // Установяване на нов примитивен елемент
        void SetAlphaVector();              // Намиране на всички примитивни елементи
        IntegerVECTOR GetAlphaVector();  // Връща всички примитивни елементи
        void SetCardinality(int newcard);   // Задаване на нова характеристика p
        int GetValueAlpha();                // Връща примитивния елемент
        virtual CString ValueString(int a);  // Извежда α^a
        virtual int GetAlphaPower(int pwr);  // Връща α^pwr
        virtual int Parse(CString str, int onempty);  // Проверява валидността на стринга
        static CString CleanString(CString str);      // Изчиства интервалите в стринг
        virtual char GetAlpha();            // Връща примитивния елемент като символ
        virtual int GetIntAlpha();          // Връща примитивния елемент като число
        static int Modulo(int a, int b);    // Връща a mod b
        virtual int Divide(int a, int b);   // Деление a / b в GF(p)
        virtual int Product(int a, int b);  // Умножение a . b в GF(p)
        virtual int Minus(int a, int b);    // Разлика a – b в GF(p)
        virtual int Sum(int a, int b);      // Сума a + b в GF(p)
        int Normalise(int Gelement);  // Нормализира отрицателните елементи
        int GetCardinality();               // Връща характеристиката p
        int GetDisplayMode();               // Връща режима на изобразяване
        void SetDisplayMode(int newDispmode); //Установява нов режим на изобразяване
        CGaloisField();                     // Конструктор по подразбиране
        CGaloisField(int p);                // Конструктор при зададена характеристика p
        CGaloisField(int newcard, int newdmode, int newalpha);
        // Конструктор при зададени p=newcard; DisplayMode=newdmode; Alpha=newalpha
        virtual ~CGaloisField();            // Виртуален деструктор
  private:
        IntegerVECTOR AlphaVector;// Вектор от всички примитивни елементи
        static boolean IsAlpha(int myalpha, int num);
                                    // Проверява дали myalpha е примитивен елемент
        static int FindAlpha(int num);      // Намира първия примитивен елемент α
        static int Prime(int p);            // Проверява p дали е просто
        int GPower(int what, int vl);  // Пресмята what^ vl в GF(p)
        int Alpha;                  // Примитивен елемент α
        int DisplayMode;            // Режим за извеждане на екрана: числова редица или полином
        int Cardinality;            // Характеристика p
};
```

На основата на теорията на пръстен от полиноми, деление и разложимост на полиноми от съвременната алгебра [1, 2, 3] е създаден ***клас*** **CGaloisPolynomial** *на* Visual C++ 6.0. *Той се характеризира със следните* свойства*: поле на Галоа* CGaloisField* Field, *степен* Degree*, променлива* char Alpha *и коефициенти* int* Coefs*[]*.

Създадените два класа позволяват генериране на полета на Галоа с характеристика $p \leq 2^{32}$, както и извършването на действията събиране, изваждане, умножение и деление с остатък на полиноми в тези полета (фиг. 2). Създадената програма *Galois Field* дава възможност за запис на получените резултати в текстов файл.



Фиг. 2. Общ изглед на приложението *Galois Field*

V. Заключение

Предложеният ускорен алгоритъм за генериране на просто поле на Галоа *GF*(*p*) многократно подобрява бързодействието като с увеличаване на *p* бързодействието се увеличава експоненциално, за сметка на намаления брой проверки с цел елиминиране на елементите, които не могат да бъдат примитивни.

Създадените програмни модели на двата класа *CGaloisField* и *CGaloisPolynomial* позволяват генериране на полета на Галоа с характеристики $p \leq 2^{32}$, както и извършване на действията събиране, изваждане, умножение и деление с остатък на полиноми в тези полета. С тези си свойства програмата *GaloisField* може да се използва:

- за обучение в областта на математическите основи на шумостойчивото кодиране и криптографията при малки стойности на характеристиката *p* на полето на Галоа;
- за изследвания и криптоанализ при големи стойности на характеристиката *p* на полето на Галоа.

ЛИТЕРАТУРА

[1] ***Ленг*** *С.,* Алгебра, Перевод с английского Е. С. Голода, Мир, М., 1968.
[2] ***Лилд*** *Р.,* Нидеррайтер Г., Конечные поля, том 1 и 2, Перевод с английского под редакции Начева В.И., Мир, М., 1988.
[3] ***Baker*** *A.,* Algebra & Number Theory, Department of Mathematics, University of Glasgow, http://www.maths.gla.ac.uk/~ajb.
[4] ***Clark*** *G. C.*, Cain J. B., Error - Correction Coding for Digital Communications, Plenum Press, New York and London, 1987.
[5] ***Das****, A.*, Algebraic Number Theory, Department of mathematics, Indian Institute of Technology, Kanpur, pp. 97, India, 2002.
[6] ***Menezes*** *A.*, P. van Oorschot, and S. Vanstone, "Handbook of Applied Cryptography", CRC Press, 1996, www.cacr.math.uwaterloo.ca/hac
[7] ***Mollin*** *R. A.*, "Algebraic number theory", CRC Press Series on Discrete Mathematics and Its Applications, Chapman & Hall, 2000.
[8] ***Schneier*** *B.* "Applied Cryptography", Jhon Wiley & Sous Inc., 1998.

НЕОБХОДИМОСТ ОТ СИСТЕМЕН ПОДХОД ПРИ ИЗГРАЖДАНЕТО НА СИСТЕМА ЗА ДИСТАНЦИОННО ОБУЧЕНИЕ

ПЕТЪР ИВ. СТОЙКОВ

NEED FOR A SYSTEMATIC APPROACH WHEN BUILDING THE SYSTEM OF REMOTE EDUCATION

PETAR IV. STOIKOV

The European tendencies, the work of the Ministry of Education and Science and particularly the work of the Remote Education Center for Humanitarian Studies from the Shumen University "Bishop Konstantin Preslavsky" are presented in short.

The author, using a systematic approach, looks at the three types of interaction in the system of remote education: functions, stages, and technologies.

KEY WORDS: eLearning education

Дистанционната форма на обучение е качествено прогресивен начин за получаване на знания, които осигуряват максимална възможност на всеки желаещ да се учи, независимо от разстояние, възраст, здравословно състояние, служебна ангажираност и други фактори. Международната организация ISTE (Internacional Society for Technology in Education), със седалище във Вашингтон, описва дистанционното обучение като "Съчетание на обучението, от което се нуждаем с удобството и гъвкавостта, които желаем".[15]

От началото на 90-те години въпросите, свързани с изграждането и развитието на информационното общество, са на прякото внимание на Европейската комисия, Съвета на Европа и Европейския парламент. През изтеклия период от време тези институции предприеха множество крупни инициативи и разработиха редица програми и планове, целящи да се активизира участието на европейския научен и бизнес потенциал в тези процеси. Най-важните от тях, според центъра за популяризиране на информационното общество, са[16]:

- **Инициативи**
  - Електронна Европа (eEurope)
  - Стани цифров (GoDigital)
  - Елекронна търговия за промишления сектор (ECOM-IS)
  - Електронно обучение (eLearning)
    hhtp://europa.eu.int/comm/education/elearning/index.html
    с бюджет: 7,5 млн. евро
- **Програми**

.....

- **Планове**

.....

У нас МОН работи активно по това направление, като има разработен доклад към меморандума на ЕК "Учение през целия живот" на тема "Програма за продължаващо обучение"[17]

[15] ISTE, International Society for tehnology in Education, http:/www.iste.org/publications/index.cfm

[16] Електронно обучение. Сборник статии и доклади, С. 2004.

[17] Доклад към меморандум на ЕК "Учение през целия живот". София, МОН, 2001,

Българското законодателство в закона за висшето образование[18] в чл.42 алинея (9) е посочено, че "Формите на обучение във висшето училище са редовни, задочни, вечерни и дистанционни.", т.е. **дистанционната** форма може да се използва за придобиване на образователната степен "специалист", "бакалавар" и "магистър".

В Шуменския университет "Епископ Константин Преславски" се реагира своевременно за внедряването на тази нова форма на обучение като бе създаден Център за дистанционно обучение по хуманитарни науки, разработен е и Правилник за същия[19]. От центъра са разработени Стандарти и препоръки за съставяне на учебници. В тях са дадени подробни дидактически указания за оформяне на същите.

Започнаха и да се водят първите учебни дисциплини по тази форма. Основната част от учебниците са в електронен вид преработени съществуващи учебници, които са допълнени със съответни тестове и други подобрения.

Това в определена степен отговаря на настоящото положение на дистанционното обучение и в другите учебни заведения, където се използват от въведени във електронен вид обикновени учебници, до модерни Web-базирани системи за дистанционно обучение, които могат да се групират в:

- Системи за управление на Web-базирани курсове CMS (Course Management System)[20]. Това към момента е най-голямото по количество системи за Web-базирано обучение. Българска система от този тип е системата eLSe на Русенския университет.[21]
- Системи за управление на Web-базирано обучение LMS (Learning Management System)[22]. Пример за типични LMS са: Docebt, Generation 21, Central и др.
- Интегрирани образователни среди – LMR платформи. Според Robert Jackson "Вероятно LMS-платформите са най-големия и най-конкурентноспособен пазарен вариант на образователните системи в днешно време, който е готов за консолидиране в близко бъдеще."[23]

Системният подход отграничава три типа взаимоотношения в системата за дистанционно обучение: функции, етапи и технологии. [24] (Сх. 1)

Функциите са: осигуряване, управление и обслужване.

Първата функция е необходима за осигуряване на учебното съдържание и обезпечаване на методите и формите от технологията на дистанционното обучение в съответствие с изискванията на дидактическата наука и необходимите средства за комуникация и взаимодействие.

Дистанционното обучение е целенасочен управляем процес, което изисква от системата да изпълнява управление на компонентите и дейностите на учебния процес за постигане на максимален образователен ефект.

---

http:/www.lifelong.learning-bulgaria.org/index.htm

18 ЗАКОН ЗА ВИСШЕТО ОБРАЗОВАНИЕ Обн., ДВ, бр. 112 от 27.12.1995 г.

19 Правилник на Центъра за дистанционно обучение по хуманитарни науки, утвърден на заседание на Факултетния съвет на Факултета по хуманитарни науки – Протокол N 2/18.10.2002

20 виж. Belyk, D.,Schubert J.&J. Baggaley. Clasification of DE delivery systems, International Review of Research in Open and Distance Learning, (IRRODL), http://cde.athabascau.ca/softeval/

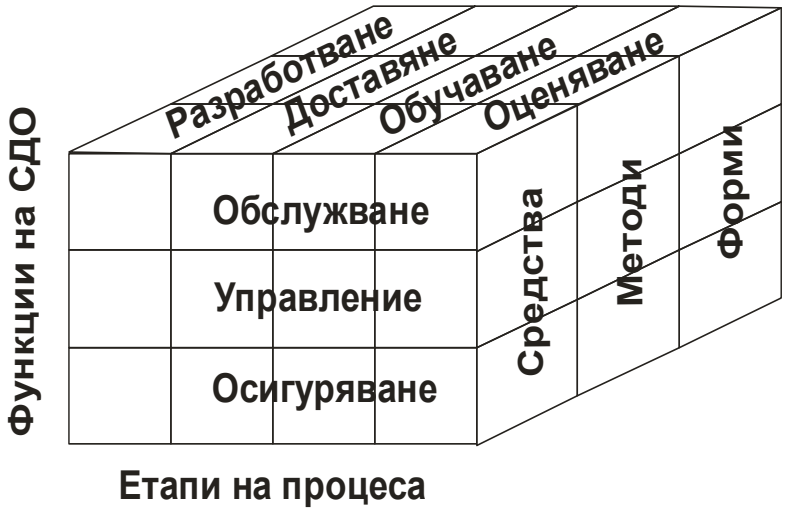
21 Hristov, T. An apprach to development of e-learning Software platform, Proceedings of the Compsys Tech'2002, Sofia, 2002

22 Integrity e-Learning, http://www.ielearning.com/wbt/index.cfm

23 Jackson R. Weblearning Resources,Categorization LMS by Function http://www.knowledgeability.biz/weblearning/#OverviewsandReferences

24 С. Ненова, Моделиране на Web-базирана система за дистанционно обучение, дисертация, С. 2004.

Трета група функции е свързана с реализиране на технологията на дистанционното обучение и обхваща дейности по подготовка, поддръжка и експлоатация на средствата за обучение и контрол, подпомагане реализацията на различните методи и форми.



Етапите на жизнения цикъл на процеса на дистанционното обучение обхващат: разработване, доставяне, обучаване и оценяване, които по-пълно звучат така:

- подготовка и разработване на образователното съдържание (Autohoring);
- комплектоване, предаване и доставяне на учебното съдържание (Delivery);
- управление на усвояването на учебното съдържание и взаимодействията между субектите (Education);
- контрол и оценка на резултатите в съответствие с поставените цели и корекция в учебното съдържание (Assessment).

По-широкото тълкуване на понятието "технология" в сферата на образованието включва начина на реализация на учебния процес, обхващайки педагогическите, организационните и техническите компоненти.[25] Технологията в този смисъл обхваща организационните форми, педагогическите методи и техническите средсва, и дава отговор на въпросите "по какъв начин" и "с какви средства" се постигат поставените образователни цели, т.е. под технология на дистанционното обучение ще разбираме съвкупност от методите, формите и средствата за реализиране на дистанционното обучение.

В процеса на дистанционното обучение намират приложение всички известни форми за обучение като: лекции, семинари, лабораторни занятия, контролни работи, курсови работи, изпити, консултации, но се прилагат с определени особености и специфика.[26]

Изследванията показват, че при дистанционното обучение са приложими всички общо-дидактически методи за обучение, които са добре известни в практиката и описани в световната литература, а именно: информационно-рецептивен, репродуктивен, проблемно ориентиран, евристически и изследователски.[27]

Според специалисти от Американския институт за изследвания в развитието на дистанционното обучение могат да се разграничат четири основни етапа, или поколения, свързани най-вече с информационните носители, използвани за трансфер на учебното

[25] А. Андреев "Дидактические основй дистанционного обучения", Москва, 2 000.
[26] Золотарев А., колектив, Теория и методика систем иненсивного обучения, Москва МГТУ ГА, 1994
[27] И. Я. Лернер, Дидактические основй методов обучения. Москва, Педагогика, 1981.

съдържание и за осъществяване на контакт между обучаващи и обучаеми.[28] Тези етапи най-общо могат да се представят по следния начин:

I) 1840-1960 – печат, радио, телевизия;

II) 1960-1985 – аудио, факс;

III) 1985-1995 – електронни и компютърни технологии;

IV) 1995-..... – интернет и Web-базирани технологии.

Всеки следващ етап добавя нови информационни носители, като неизключва употребата на съществуващите до момента.

Един от общопризнатите специалисти в дистанционното обучение Moor, изтъква, че за ефективно дистанционно обучение е необходимо максимизиране и на трите основни типа взаимодействие[29]:

- обучаем – учебно съдържание;
- обучаем – преподавател;
- обучаем –други обучаеми.

Това изисква осигуряване на технологични средства за реализиране на всяко от тези взаимодействия.

Съществуват различни средства за взаимодействието "обучаем – учебно съдържание" от обикновените печатни материали, популярните аудио и видио средства, телевизия, радио, преносимите компютърни средства (дискети и CD) и се стигне до съвременните Internet и Web технологии.

Интернет създаде ново поколение дистанционно обучение (четвърто поколение) въвеждайки интелигентна комуникационна среда и създавайки преносим модел за широко участие в проектирането и използването на курсове за дистанционно обучение.[30]

Средствата за комуникация между субектите на дистанционното обучение са основно Интернет и се делят на две основни групи:

- Интернет технологии за синхронна комуникация: Chat IRS (Internet Relay Chat), MOO (Multi-user Object Orient) – това са виртуални стаи (virtual rooms), MUD (Multi User Domain), White board (интерактивна виртуална среда) и компютърна видео конференция.

- Интернет технологии за асинхронна комуникация: e-mail, USENET, Telnet.

Според препоръките на Генералната дирекция по образование и култура към Комисията на Европейския съюз един от основните критерии, по които ще се оценява степента на използване на съвременните информационни и комуникационни технологии в учебния процес на всеки университет, ще бъде броят на WEB базирани курсове, т.е. броят на дисциплините, за които са създадени WEB сайтове.[31] Един от начините за бързо изпълнение на изискването е всеки преподавател да направи учебните си материали достъпни чрез Интернет е използването на т. нар. софтуерни платформи[32] за електронно обучение, които представляват интегрирани среди за разработване, модифициране и администриране на WEB базирани учебни пособия и за комуникация между преподаватели и студенти, което бе разгледано в първата половина на публикацията.

---

[28] L. Douglas, American Institutes for Research, Report of National Celter for Education Statistics, 1998

[29] Moor, M.G., Theory of transactional distance, in Keegan, D., Theoretical Princiles of Distance Education, Routledge, London, 1993

[30] Passerini, M.K., Granger, J., A developmental model for distance learning using the Internet., Computer&Education, 2000

[31] Смрикаров, А., А. Василева Инициативата "Електронно обучение" на европейската комисия. Автоматика и информатика, С., 2002

[32] Христив, Ц., колектив, Един подход към създаването на софтуерна платформа за електронно обучение, Електронно обучение. Сборник статии и доклади, С. 2004.

Обощавайки всичко до тук, може да се каже, че една софтуерна платформа за електронно обучение трябва да отговаря на следното:

- Да позволява създаванетона WEB бзирани курсове, отговарящи на изискванията на Генералната дирекция по образование и култура към Европейската комисия[33], т.е. да дава възможност за публикуване, допълване и коригиране на материали, съдържащи:
  - анотация на дисциплината;
  - учебна програма;
  - литература;
  - лекции;
  - упражнения;
  - тестове;
  - задачи;
  - график на занятията;
  - съобщения за контролни и изпити;
  - конспект за изпита;
  - информация за преподавателския екип;
  - форум за обсъждане на въпроси, поставяни както от преподавателя, така и от студентите.
- Да включва средства за водене на статистика на курса и по-точно регистриране на посещения в сайта като цяло и в частност на отделни учебни единици от отделните студенти, на успеха на студентите, на степента на усвояване на различни теми от лекционния материал и т.н.;
- Да бъде универсална и същевременно да позволява приспособяване към структура и изисквания на конкретен университет;
- Да не изисква специални познания и умения по Интернет програмиране и WEB дизайн от страна на преподавателите;
- Да не изисква големи системни ресурси от страна на сървера;
- Да не изисква предварително инсталиране на допълнително програмно осигуряване на потребителските станции освен стандартен WEB браузър;
- Да е съвместима с най-разпространените операционни системи и WEB браузери;
- Да предоставя възможност за осъществяване на връзка между студентите и преподавателския екип;
- Да позволява лесна смяна на езиците;
- Да е защитен от неоторизиран достъп;
- При възможност да бъде направена с използането на безплатни софтуери, за да бъде по-евтина и др.

[33] Смрикаров, А., А. Василева Инициативата “Електронно обучение” на европейската комисия. Автоматика и информатика, С., 2002

ЕДИН ПОДХОД ЗА АНАЛИЗ НА ПРЕДМЕТНА ОБЛАСТ НА ИНТЕЛИГЕНТНА СИСТЕМА ЗА УПРАВЛЕНИЕ НА ОБУЧЕНИЕТО

НАЙДЕН В. НЕНКОВ

AN ANALYSIS APPROACH AT OBEJECT AREA FOR UNIVERSITY TREANING MANAGING SYSTEM

NAYDEN V. NENKOV

This article offer ontological approach for a conceptual knowledge analysis for specified object area – university teaching subject managing. It's applicable to solve problems knowledge conceptualization and formalization at the decision support system for university teaching intelligence system creating.

KEY WORDS: Object area, Conception, Knowledge, Ontology, Ontology development tool

УВОД

Управлението на учебната дейност е сложна оптимизационна задача, която се зависи от множество фактори. Това често е свързано както с рутинни действия на преподавателите по обслужване на информационния поток, така и с техни експертни решения. Един ефективен подход за решаване на съществуващите проблеми е създаването на интелигентна система за управление на обучението, базираща се на експертни знания.

Известни са множество решения, предлагащи автоматизирани курсове по отделни дисциплини или системи за цялостното обучение по e-learning технологията.

Системата Diogene [2] е Web базирана независима платформа за обучение на специалисти по информационни и комуникационни технологии. Предназначена е за подпомагане на обучаемите през целия цикъл на подготовка, от дефинирането на целите, до оценяването на резултатите, чрез индивидуализирани, адаптиращи се към нуждите на обучаемите курсове. Тя използва съвременни технологии като: мета-данни и онтологии за обработване на данни, моделиране на обучаемите, гъвкаво адаптиране на обучението към нуждите на обучаемия, кооперативно групово обучение и он-лайн подпомагане. Предлага и множество от иновативни характеристики, като: динамични стратегии за учене, семантична отвореност на WEB технологията, базирана на Semantic Web и XML.

В рамките на изградения виртуален университет [3] у нас се предлагат Web базирани курсове на обучение чрез софтуерната платформа "e-Learning Shell". Тя позволява на преподавателите да добавят учебни планове, курсове за обучение, анотации, теми на занятията, библиография, лабораторни упражнения, актуални проблеми, конспекти и други необходими учебни материали.

Широко разпространение у нас и в чужбина е получила Moodle [4] система за управление на курсове (course management system - CMS). Тя предлага богат набор от инструменти за разработване и поддръжка на WEB–базирано обучение. Разпространява се безплатно по инициативата Open Source, но е надеждна и удобна за използване.

Повечето от известните разработки в тази област, успешно се прилагат при създаването на обучаващи системи по досегашните технологии, но не и при системи използващи знания.

Анализирането на структурата на предметната област, същността на обектите и съществуващите отношения между тях е важен етап от реализирането на интелигентната система. Все по-често, това се извършва чрез средствата на така наречения „онтологичен

инженеринг” [5], а разработването на онтология е ефективен начин за разкриване на важните свойства и отношения в предметната област на интелигентната система.

Целта на статията е да покаже един подход за изграждане на онтология на предметната област – университетска структура за обучение, която да се ползва при концептуализация и формализацията на знания за управление на обучението по специалност.

Моделът на областта е разработен чрез среда за разработване на онтологии Protégé [5] и UML диаграми [1].

ИЗЛОЖЕНИЕ

Интелигентната система цели да повиши ефективността при управление на учебната дейност на преподавателите от университетите и висшите училища. Тя улеснява използването на съществуващата базата данни за учебния процес и съветва за необходимите управленски решения, които трябва да се вземат.

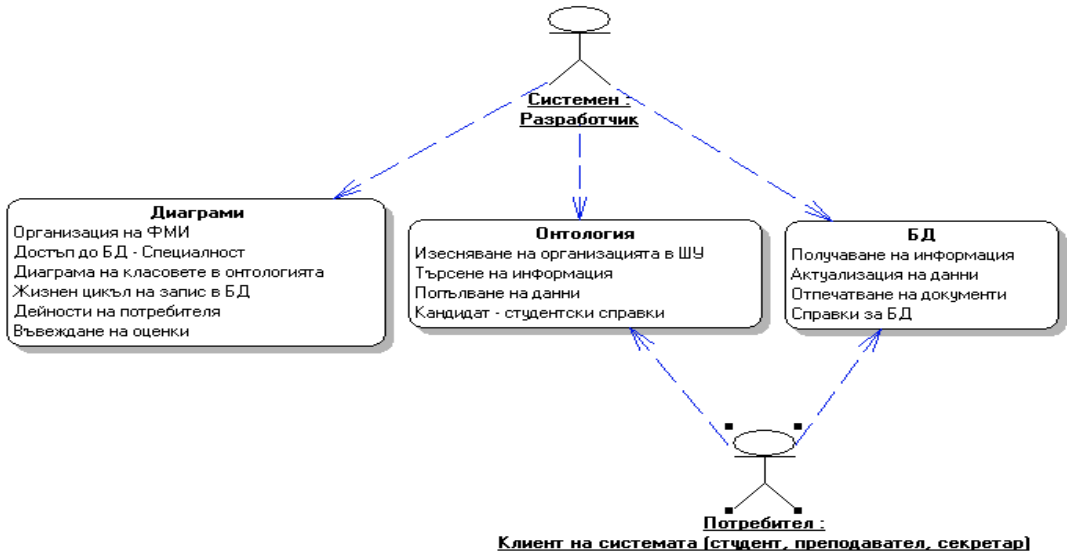
В статията се разглежда онтологичен подход за анализ на система за управление обучението във Факултета по математика и информатика на Шуменския университет по специалността Информатика - бакалаври. Възможно е неговото прилагане и за други специалности, и в други университети.

Концептуалното представяне на знания представлява описание на множество обекти и понятия, знания за тях, връзки между тях, атрибути на понятията (слотове) и ограничения, наложени от слотовете (фасети). Онтологията се състои от термини, организирани в таксономични речници, техните определения и свойства, а също и свързаните с тях аксиоми и правила за извод [5].

В центъра на онтологията се намират класовете, с които се описват понятията от предметната област. Възможно е те да имат подкласове, с които се представят по конкретни знания, отколкото класовете на по горното ниво (суперкласовете). Всеки представител на класа се нарича екземпляр. Слотовете описват свойствата на класовете и екземплярите. Всички екземпляри от определен клас имат слот, чието значение се явява екземпляр на класа.

Разработката на онтология включва: определяне на класовете в онтологията; разположение на класовете в таксономична йерархия; определяне на слотовете и описание на допустимите им значения; запълване значението на слотовете екземпляри.

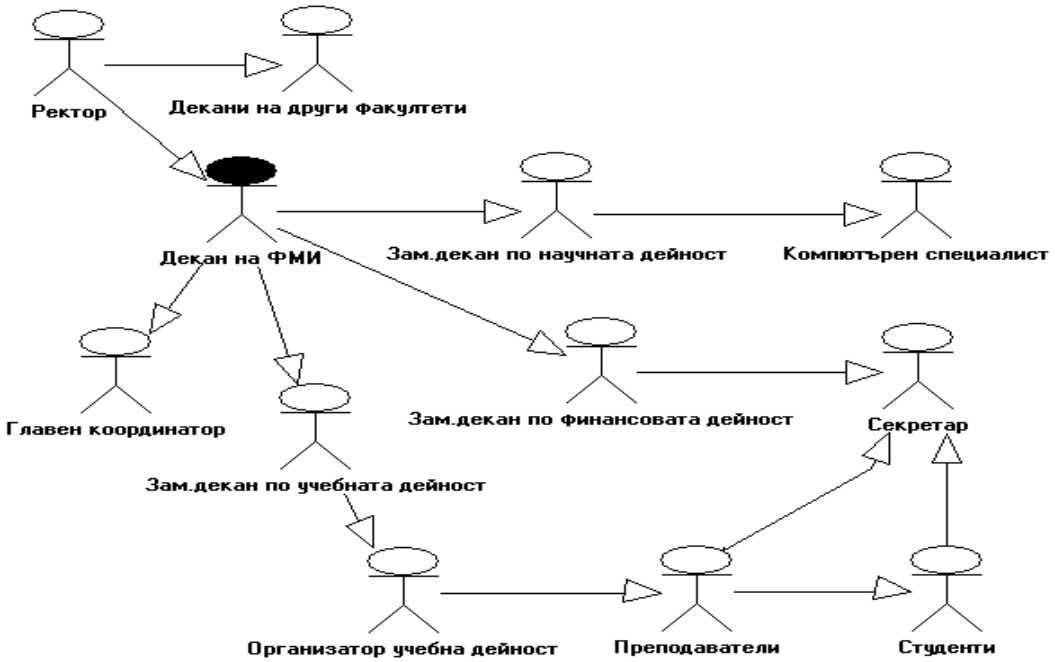
След това може да се създаде база знания, да се определят отделните екземпляри на съществуващите класове, да се въведат значения и допълнителни ограничения за слота и да се създадат отношения между екземплярите (фиг.1).



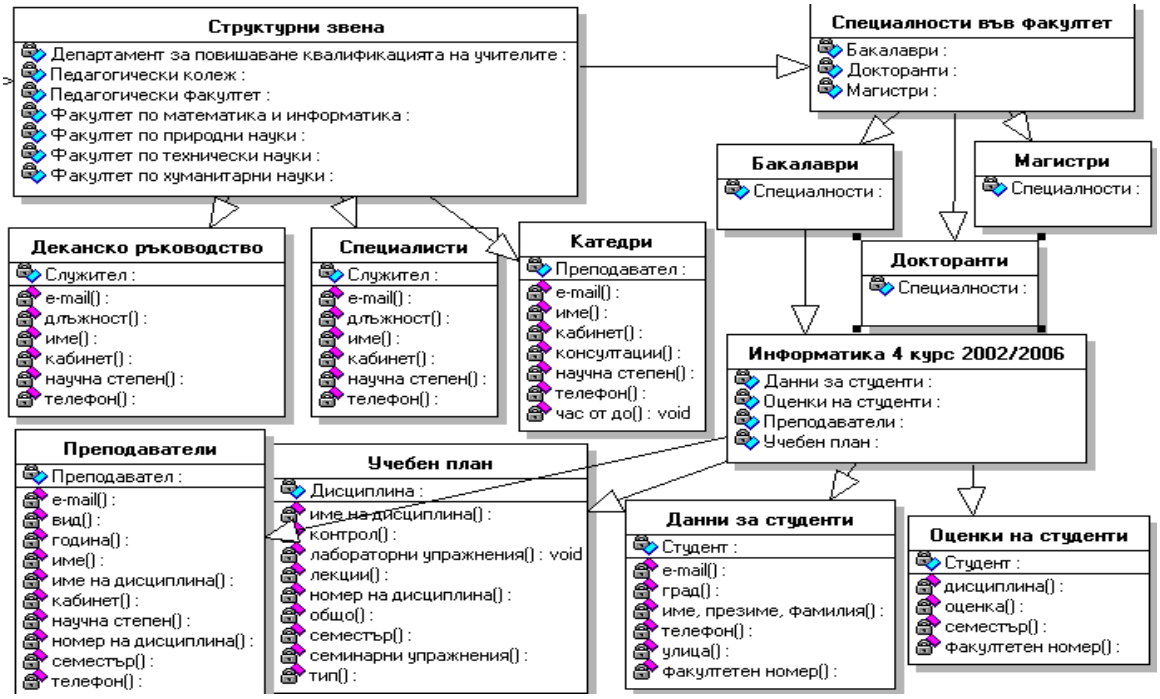
Фиг. 1. Реализация на интелигентна система за управление на обучението

1. Модел на системата за управление на обучението

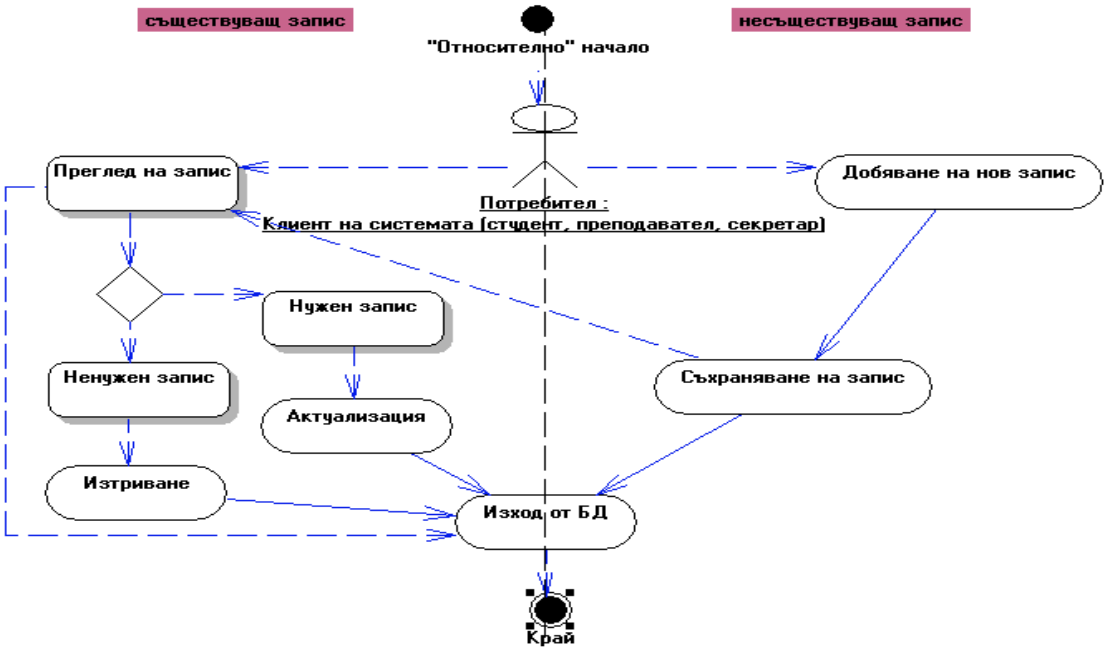
Моделът на системата е отговаря на описаната в Правилника на университета структурата и организацията на учебния процес. Чрез него се анализират информационните потоци и събитията свързани с управлението на учебната дейност. Графичното му представяне са Use-Case диаграмите [1], които са показни на фигурите - 2, 3 и 4.



Фиг.2. Диаграма на взаимодействията



Фиг. 3. Диаграма на основите елементи на обучението във факултета



Фиг. 4. **Жизнен цикъл на запис в БД**

2. Създаване на онтология на предметната област

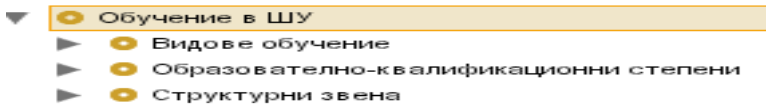
Онтологията описва структурата на отделните звена в Шуменски университет. По-конкретно е представена структурата и информационните връзки във Факултета по математика и информатика и специалността „Компютърна информатика". В онтологията са показани само университетската структура и една специалност. Онтологията може да се актуализира, като се добавят нови класове, подкласове, слотове, фасети и състояния (екземпляри). Така тя може да придобие огромен размер, но това е извън рамките на статията. По-нататък са описани основните и елементи.

2.1. Класове и подкласове

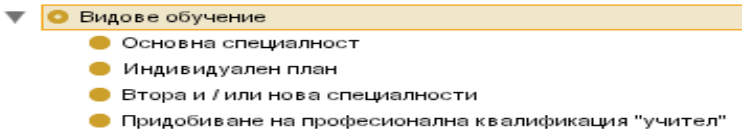
Класовете в онтологията изясняват понятията, които се включват в предметната област обучението в ШУ. Класовете се задават в секцията - ***Classes***.

Класовете в онтологията може да бъдат конкретни (***Concrete***) и абстрактни (***Abstract***). Когато един клас е ***Concrete***, тогава е възможно да се създават негови екземпляри. По принцип на класовете, които са поднива в йерархията се поставят в роля ***Concrete***, а всички останали класове в роля ***Abstract***.

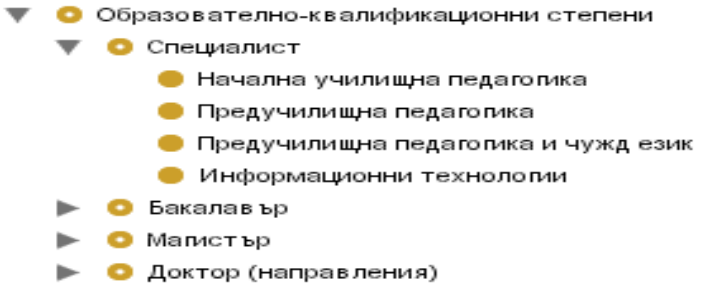
Главният клас "Обучение в ШУ" (фиг. 5) се разделя на 3 подкласа, с които се описват различните видове обучения (фиг. 6) в университета, образователно-квалификационните степени (фиг. 7) и структурата на университета (фиг. 8).



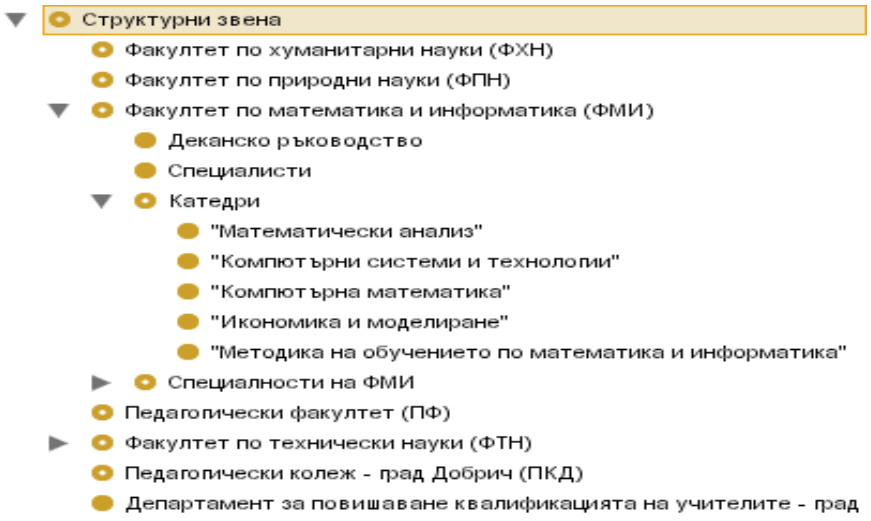
Фиг. 5. **Структура на класа "Обучение в ШУ"**



Фиг. 6. **Подклас "Видове обучение"**

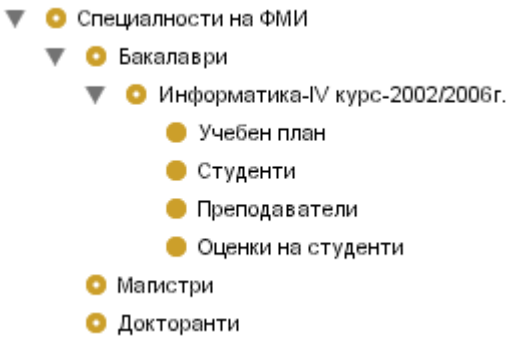


Фиг. 7. **Подклас "Образователно-квалификационни степени"**



Фиг. 8. **Подклас "Структурни звена"**

В подкласа на "Специалности във ФМИ" (фиг. 9) е описана подробна информация за специалността – Информатика 4 курс, редовно обучение с нейните студенти, преподаватели, учебен план и оценки на студентите.



Фиг. 9. **Подклас " Специалности във ФМИ "**

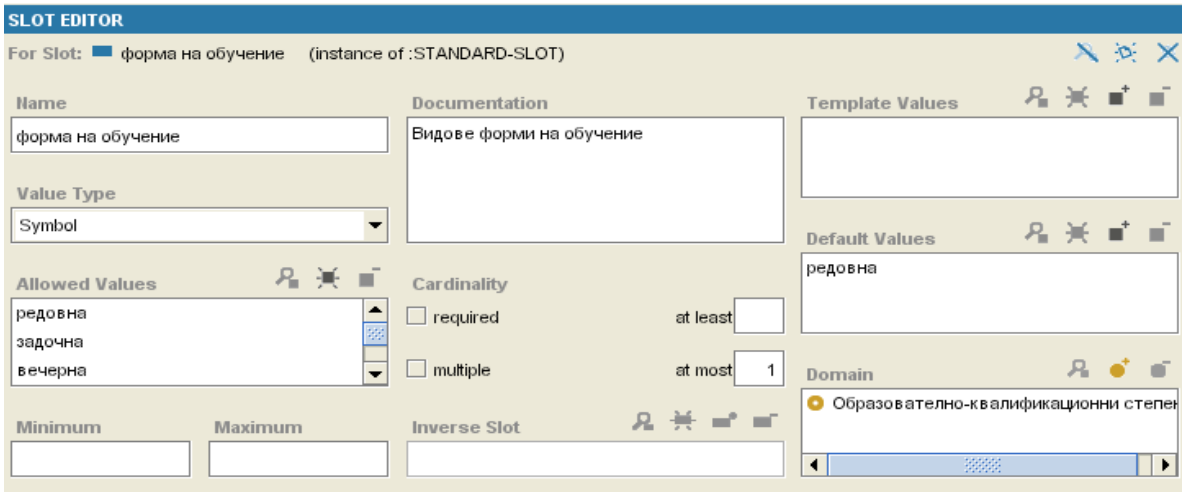
За някои подкласовете в цялата йерархия, които имат роля ***Concrete*** се създават екземпляри. Екземплярите се явяват листата в дървовидната структура, но те се реализират след като се определят слотовете и формите.

2.2. Слотове и фасети

След като се създаде йерархията на класовете, за всеки клас или определени класове с роля ***Concrete*** се задава слот, който представя определено свойство на класа/класовете от секцията ***Slots***. Всички слотове приложени за даден клас определят множеството от допустими му стойности, а самият клас се нарича ***домейн***. Фасетите представляват ограничения върху множеството от допустими стойности и са вложени в слотовете.

Слотовете се прилагат за даден клас или класове от обекти, като всички подкласове и поднива наследяват свойствата на суперкласа си, освен ако не бъдат променени или скрити.

В зависимост от типа на слотовете се определят как да изглеждат формите – като падащи списъци, текстова кутия, текстова област, със стойности по подразбиране, кутия с отметка и др. Пример за слот – "форма на обучение" е показан на фиг. 10.



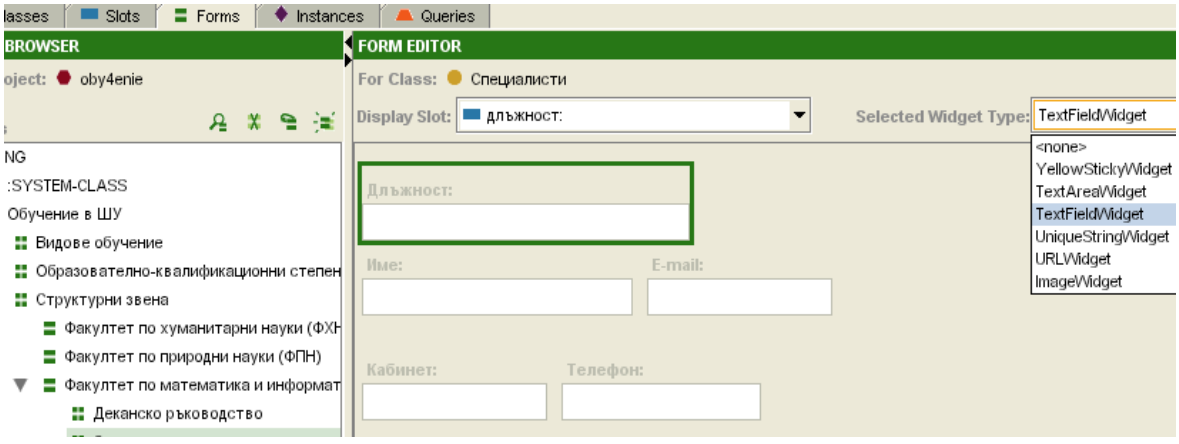
Фиг. 10. **Слот " Форма на обучение "**

В областта ***Cardinality*** се определя броя на екземплярите на класа, за които е приложен слота - домейн. В случая за класа "Образователно-квалификационни степени".

2.3. Форми

Не е необходимо да се създават форми, по подразбиране Protégé ги създава при изграждане на класовете. При отваряне на секцията ***Forms***,определяната форма може да се приложи за клас с подкласове или краен клас, като формата се наследява и от всички подкласове, но това зависи и от ролите на класа "родител" и класовете "деца". Свойства (слотовете) на форма за клас могат да бъдат скрити или от различен тип. Примерна - форма "Специалисти", виж фиг. 2.16.

В падащият списък ***Display Slot*** се избира този слот, по които екземплярите на класа ще бъдат именувани. Отделните полета - слотове за дадена форма могат да бъдат разместваи по желание на разработчика на онтологията.

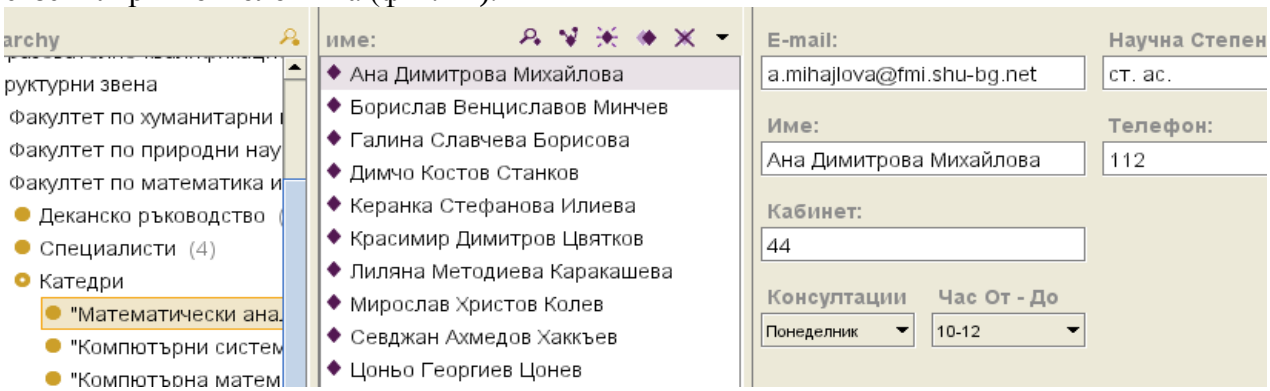


Фиг. 11. **Форма " Специалисти "**

2.4. Екземпляри

След като се създаде йерархията на класовете и се определят техните слотове и форми трябва да пристъпим към създаване на екземплярите на класовете (***instances***).

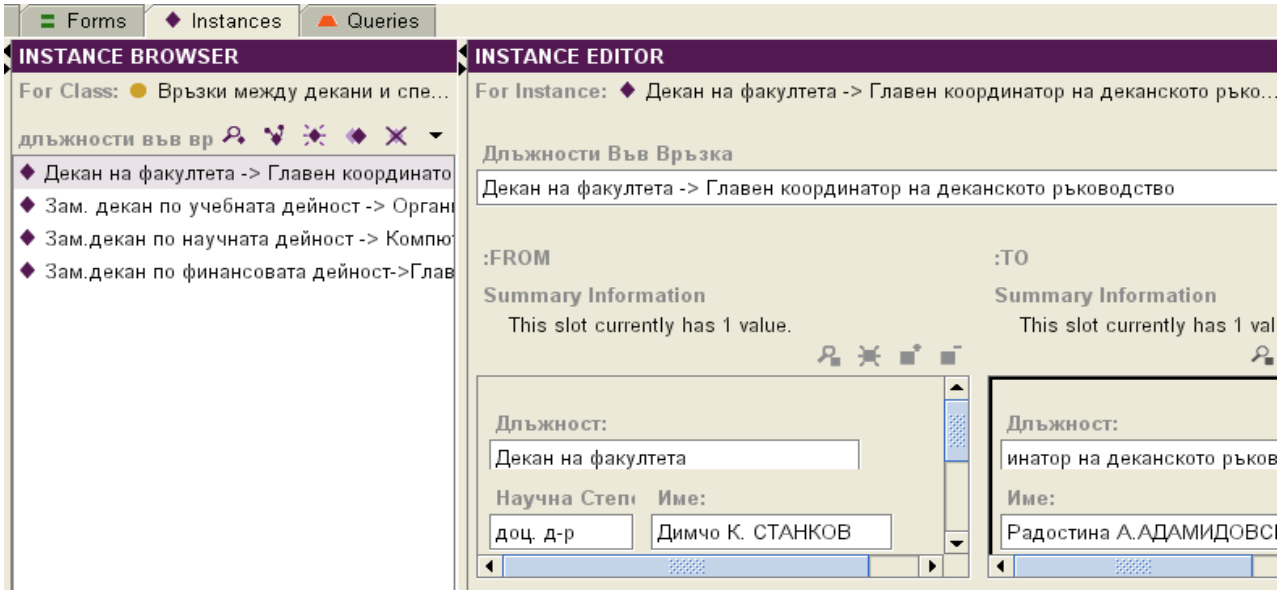
Екземплярите се явяват поднива в йерархията. Имената на екземплярите на класа могат да се повтарят, но стойността на полета-слотове в тях трябва да бъде различна. Ето пример за екземпляри в онтологията (фиг. 12).



Фиг. 12. **Екземпляр на подкласа “ Математически анализ ” на класа “ Катедри”**

2.5. Създаване на отношение между екземпляри на два класа

Тази операция има за цел да покаже релациите между обетите в предметната област. Например как би могла да се създаде връзка между декана и специалистите във факултета. Създаването на връзки между екземплярите на два класа се осъществява в системния клас ***RELATION***, а екземплярите на тази връзка се създават в секцията ***Instances***. (фиг. 13).



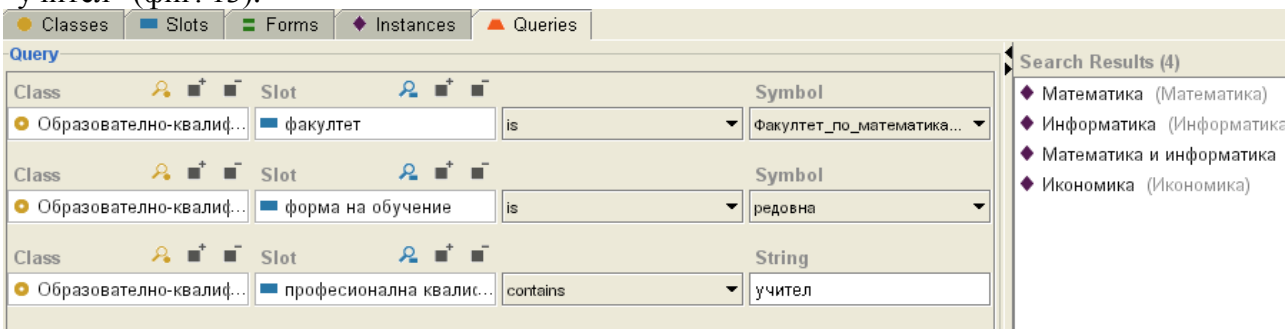
Фиг. 12. **Отношение между Декан на факултета и координатор на деканското ръководство**

2.6. Заявки

Заявките се създават с цел да се търси определена информация в екземплярите на онтологията. С помощта на заявките се прави избор на данни, удовлетворяващи дадено/и условие/я в екземплярите на класовете. Използват се операциите логическо “***и***” и “***или***”. Създаването на заявка става в секцията ***Queries*** чрез избор на клас, слот и търсена стойност. Заявките се изпълняват по време на тяхното създаване и се съхраняват в библиотека със заявки.

Резултатите от направената заявката се извеждат в прозореца ***Search Results***. За да се повтори изпълнението на дадена заявка, тя трябва да бъде отново създадена и стартирана с бутона ***Find***.

Примерът показва заявка, която извежда тези специалности от ФМИ, които имат редовна форма на обучение и предлагат на студентите професионална квалификация – “учител” (фиг. 13).



Фиг. 13. **Заявка в онтологията “ Обучение в ШУ ”**

ЗАКЛЮЧЕНИЕ

Моделите на предметната област са графични и са построени чрез разбираем език (UML). Те позволяват предварително да се анализират основните обекти и концепции.

Създадената онтология описа нагледно структурата и връзките между обектите на предметната област – обучение в ШУ, като използва йерархия от класове на фреймовия модел.

Чрез нея се извършва предварителен анализ и оценка на възможностите за изграждане на интелигентна система за управление на учебния процес по специалност.

Показани са основните етапи за изграждане на онтология, които могат да се следват и при други предметни области.

Онтологията осигурява ефективната концептуализация и формализация на данните и знанията, които са необходими за разработване на базата от данни и знания на интелектуалната система.

Тя позволява да се верифицират и отстранят евентуалните концептуални и логически грешки в модела на знанията и данните.

Разработените структурни модели и онтологията са успешно приложени в разработената интелигентната система за управление на университетска специалност.

ЛИТЕРАТУРА

1. Стоянов Ст. и колектив. Ръководство по софтуерни технологии. Пловдив, УИ, 2003.
2. FP IST Project "DIOGENE: A Training Web Broker for ICT Professionals" - http://www.diogene.org/
3. http://ecet.ecs.ru.acad.bg/cst04/
4. http://moodle.org
5. http:// protégé.stanford.edu

ЕДИН ПОДХОД ЗА ИЗПОЛЗВАНЕ НА ESTA ПРИ СЪЗДАВАНЕ НА ЕКСПЕРТНИ СИСТЕМИ С УЧЕБНА ЦЕЛ

НАЙДЕН В. НЕНКОВ, БОРИСЛАВ П. СТОЯНОВ

AN APPROACH TO USING ESTA FOR EXPERT SYSTEM DESIGNIG OF TEACHING GOAL

NAYDEN V. NENKOV, BORISLAV P. STOYANOV

This article offered one approach for developing expert system for learning goals. ESTA is the very comfortable tools for this action. It's providing the various things to teach the students in Expert system design.

KEY WORDS: expert system for text animation, expert system shell, expert system, section, section tree, knowledge, advice, parameter

УВОД

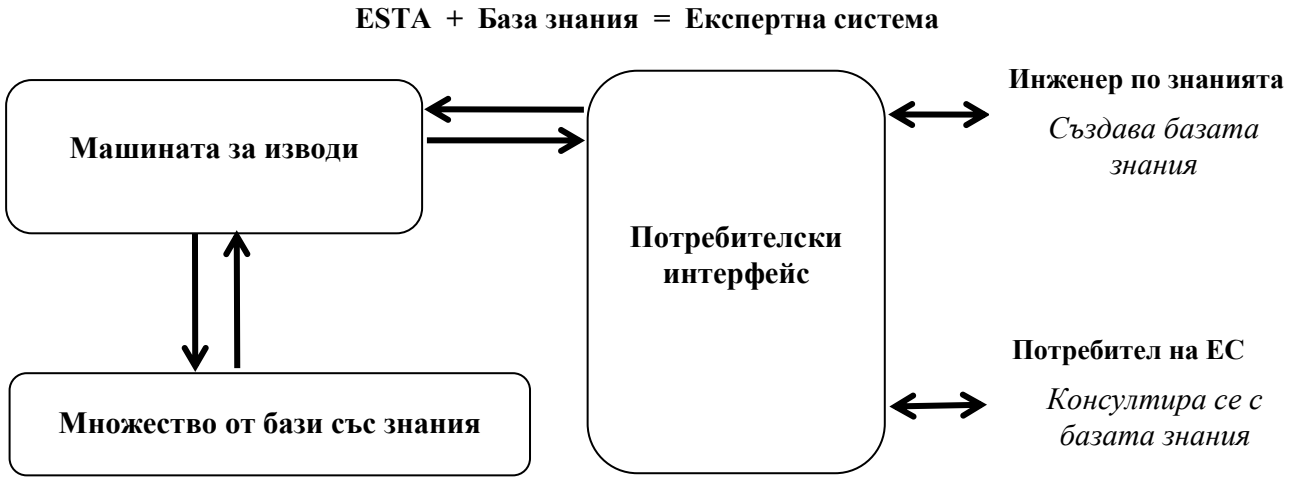
Учебната дисциплина „Експертни системи" преподава в различните университети и висши училища като задължителна или свободно избираема. Тя въвежда студентите в една от най-развитата технология на Изкуствения интелект. Подготовката на специалисти с умения за разработване на такива системи е важна задача на обучението в тази област. Често това става, като се използват традиционните софтуерни среди: Lisp, Prolog C++, Java или специализираните шел-средите: GURU, CLIPS, OPS5, KEE и други . Този подход имат много положителни страни, т.к са широко разпространени в практиката и има подготвени специалисти. Трудностите за обучаемите, произтичат от непригодността на първата посочена група средства за разработване на модел и архитектура на експертната система през етапа на проектирането. Специализираните среди са твърде сложни и изискват повече време и усилия за усвояване [1].

Затова изборът на подходящи средства за проектиране и разработване на експертни системи е важен проблем на обучението. Необходимо е чрез своите функционални възможности да улесняват представяне на знанията и проектиране на отделните компоненти от архитектурата им.. В доклада се предлага едно възможно решение, което е приложено в практиката на обучението по дисциплината „Експертни системи" с бакалавърската степен на специалностите: „Компютърни информационни технологии" и „Компютърна информатика".

ИЗЛОЖЕНИЕ

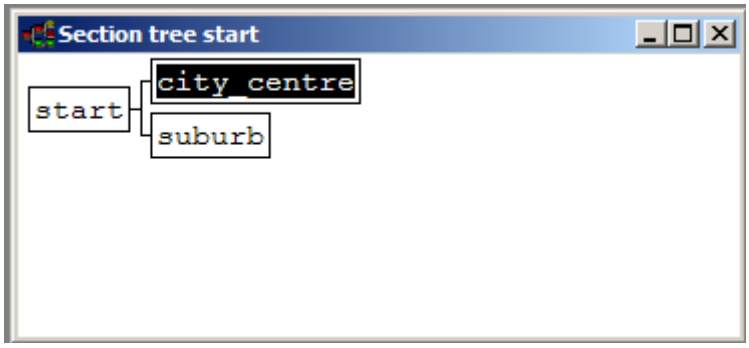
ESTA - Expert System for Text Animation е шел-среда за построяване на ЕС на PDC (Prolog Development Center), Дания. Тя е част от софтуерната среда VISUAL PROLOG и се предлага като свободно разпространявана версия, а е възможно да бъде изпробвана и on-line [2].

Характеризира се с удобен, опростен интерфейс и притежава необходимия набор средства за бързо разработване както на прототипи, така и на завършени ЕС. Със средата се предоставят множество от примерни на разработени системи и богата помощна информация. Това улеснява обучение и бързото усвояване на методологията за нейното използване. ESTA позволява модулно построяване на базите със знания за различни области и улесненото им обединяване в цялостна ЕС. Не са необходими усилия и време за разработва машината за изводи, а се използва вградената в ESTA (фиг.1). ESTA притежава удобни средства, които улесняват писането на правилата, построяването на базата знания и разработването на компонентите от интерфейса, които са стандартни за приложенията под MS WINDOS – диалогови прозорци, бутони и други. Автоматизирано е и разработването на обяснителния компонент чрез задаване на съответните текстове при отговор на въпросите "Как?" и "Защо?". Структурата и общата схема за използване на системата е показана на фиг. 1. и се базира на формулата:



Фиг. 1. Създаване на ЕС чрез ESTA

Средата е подходяща за създаване на ЕС в различни предметни области. ESTA има интерфейс към VISUAL PROLOG., който позволява интегриране с негови приложения. БЗ се изгражда като йерархична структура, за която се генерира автоматично графично изображение (Фиг. 2.). Възможно е директно писане на код и редактиране в неговите елементи.



Фиг. 2. Дървовидна схема на системата

1. **Елементи на ESTA**

Знанията се представят в три типа йерархични структури: секции (**sections**), параметри (**parameters**) и заглавия (**title**).

1.1. Секции

Синтаксисът на секцията е следния:

< секция > ::= section < име_на _секцията > [:] < текст_описание >
< списък_от_параграфи>
< списък_от_параграфи > ::= < параграф> [<списък_от_параграфи >]
< параграф > ::= [if <булев_израз>] <действие>
[if < булев_израз >] (< действие >)
< действия> ::= < действие > [,<действия >]
< действие > ::= advice | assign | call | chain | do | do_section_of | exit | stop

Име_на_секцията и **текст_описание** показват предназначението на секцията.

Параграфът е такъв тип структура на знанията, който се изпълнява в зависимост от някакво **if** условие. **Действията** могат да бъдат:

– съвет - **advice**;
– задаване на стойност за променлива - **assign**;

- извиквания на външни процедури - **call**;
- връзка с други БЗ - **chain**;
- **Пример** безусловен преход към друга секция - **do**;
- условен преход към други секции - **do_section_of**;
- изход от консултацията - **exit**;
- или временното и прекъсване - **stop**.

Всяка разработена система трябва да съдържа поне една секция с име **start**, с която започва работата на системата. Понататък се използват структурни елементи от предоставените с ESTA примерни системи: **cinema** - елементарна "съветваща" система, по какъв начин да отидем на кино; **hodge** - система за диагностика болестта на Хочкинс и **car** - системата за откриване на неизправности в автомобили (файловете с БЗ са: **cinema.kb**, **hodge.kb, car.kb)**.

Пример 1 за определяне на местоположението на киното.

section start : „определя къде се намира киното"
if (cinema = 'odeon') **do** city_centre
if (cinema = 'palace') **do** suburb

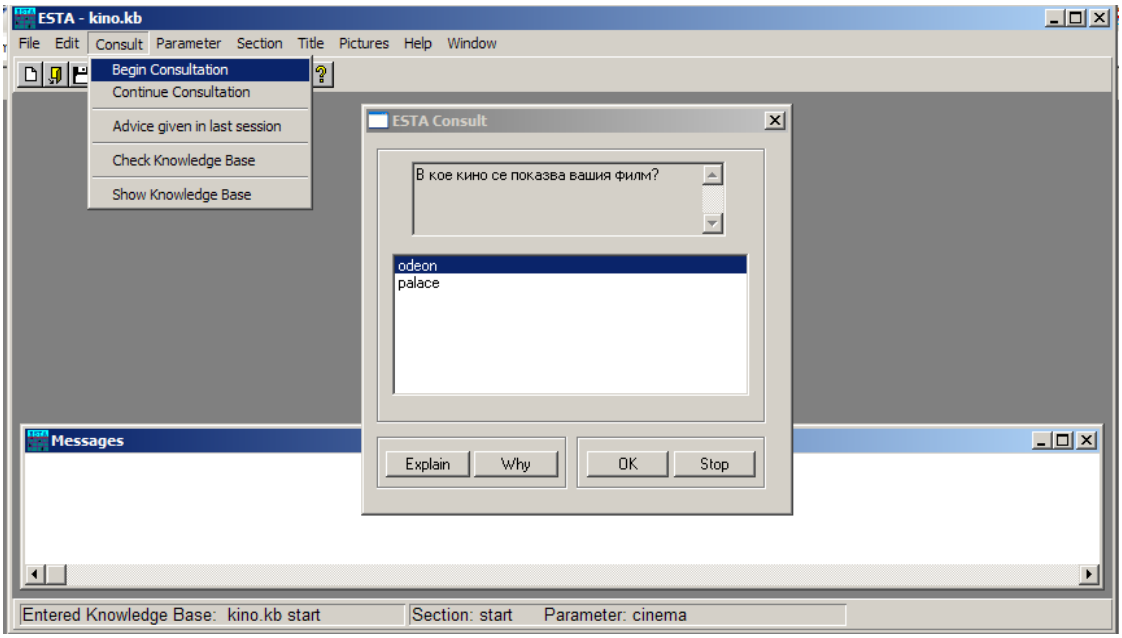
На фиг.3. е показан диалога, който генерира началната секция - **start** на системата.

1.2. Параметри

Те са 4 типа - ***boolean, text, number и category***, и се използват за дефиниране на променливи, които контролират потока от информация между секциите.

1.2.1. Параметър *boolean*

Булевите или логическите параметри се използват, когато параметрите са ограничени до една стойностите: истинска - "да", лъжлива - "не" или неизвестна - "не знам". По подразбиране ESTA, автоматично генерира диалог с възможни стойности: истина - "**Yes**", лъжа - "**No**" и неизвестна - "**Unknown**".



Фиг. 3. Начална секция **start** в ESTA

Синтаксисът за параметъра ***boolean*** е следния:

<параметър boolean > ::=
< поле за деклариране>
type boolean

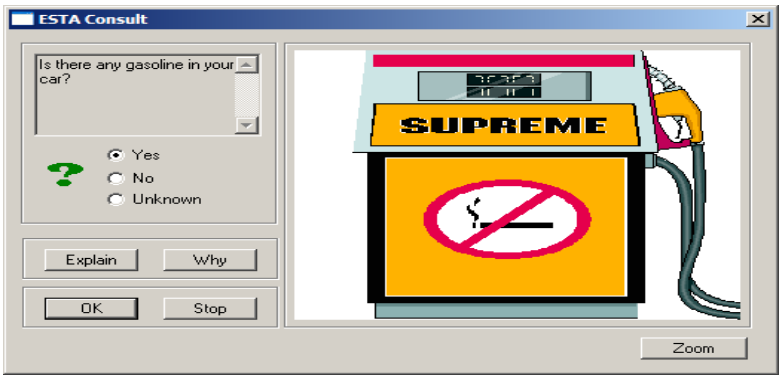
[<поле за обяснение>]
[<поле с правила (с булеви изрази)>]
[<поле за въпрос>]
[<поле за илюстрация>]

където**: поле за деклариране** – поле за описание предназначението на параметъра, **поле за обяснение** - незадължително поле, използвано при поискване на обяснение от системата, **поле с правила** - незадължителни полета за описване на правила за определяне на стойността на параметъра, **поле за въпрос** - незадължително поле за описание на въпрос и **поле за илюстрация** – незадължително поле за задаване на картинка за илюстриране на параметъра.

Пример 2: от системата за откриване на неизправности в автомобили: **car.kb.**

parameter gasoline_ok : 'there is gasoline in the car – има гориво в колата'
type boolean
explanation
'Turn on the ignition and look at the fuel gauge – Изгасете двигателя и проверете нивото на горивото'
question
'Is there any gasoline in your ' car '?' – Има ли гориво в колата ви?
picture 'gasoline'

След като се дефинира по този начин параметъра: **gasoline_ok,** автоматично се генерира следния диалогов прозорец - Фиг. 4. Илюстрацията с име 'gasoline' е зададена предварително в ESTA, чрез командата: „**pictures**"→ „**Picture database**"→ „**import**".



Фиг. 4. Диалог от параметър тип boolean

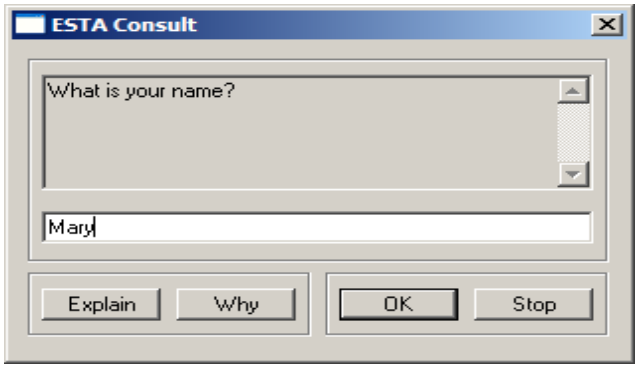
1.2.2. Параметър *text*

Текстовите параметри се използват за текстови обекти като имена на хора, любими цветове и др. Имат следния синтаксис:

<текстов параметър> ::= < поле за деклариране >
type text
[<поле за обяснение >]
[<поле с правила (с текстови изрази)>]
[<поле за въпрос >]
[<поле за илюстрация >]

Ако текстовия параметър получава своята стойност чрез отговор на зададен въпрос, ESTA ще покаже диалогова кутия с въпроса и ред за редактиране на отговора.

Пример 3: от система за диагностика болестта на Хочкинс **hodge.kb**
parameter name 'the name of the user - името на потребителя'
type text
question 'What is your name ? - Какво е вашето име?' Генерира се диалога, показан на Фиг. 5.



Фиг. 5. Диалог от параметър тип text

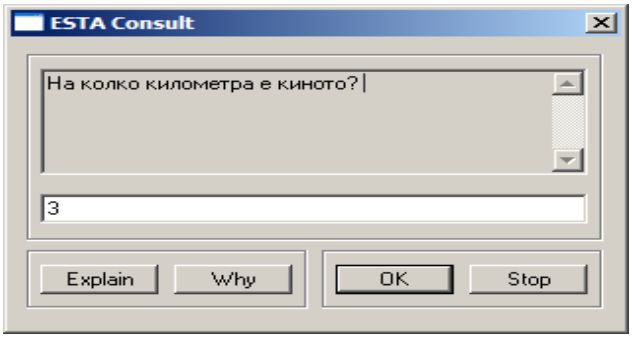
1.2.3. Параметър *number*

Синтаксисът на параметъра е:

<числов параметър> ::= < поле за деклариране >
type number
[<поле за обяснение >]
[<поле с правила (с числови изрази)>]
<поле за гранични стойности>
[<поле за въпрос >]
[<поле за илюстрация >]

< поле за гранични стойности > ::= range <минимално число> <максимално число>
поле за гранични стойности – поле, задаващо границите на стойностите на параметъра с минимално и максимално число.

Пример 4: от **cinema.kb**
parameter distance: 'отдалеченост на киното'
type number
question 'На колко километра е киното?'
На фиг. 6 е показан диалогът за параметъра distace.



Фиг. 6. Диалог от параметър тип number

1.2.4. Параметър category

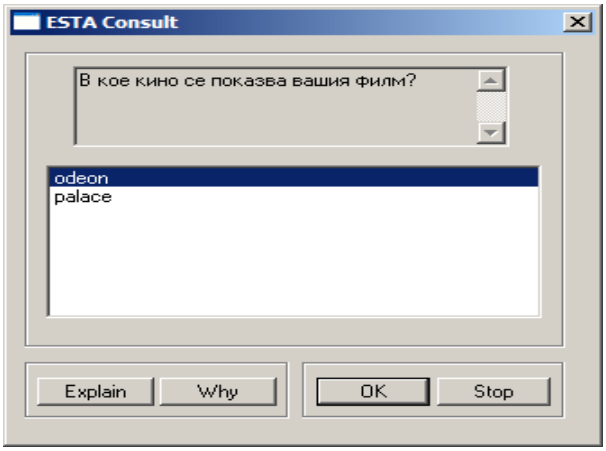
Използва се, когато получава стойност от предварително дефинирано множество стойности. Например, трябва задължително да бъде избран един от цветовете: червен, зелен, син или виолетов. Възможните стойност се описват в полето на опциите и се разделят със запетаи, а накрая - с точка. Възможно е да се зададат и обяснителни текстове. ESTA преобразува всички текстове в малки букви дори и да ползвате главни.

Синтаксисът на параметъра е:

<параметър категория> ::= < поле за деклариране >
type category
[<поле за обяснение >]
<поле с опции>
[<поле с правила (с текстови изрази)>]
[<поле за въпрос >]
[<поле за илюстрация >]
< поле с опции > ::= options <име> [- <низ>] {,<име> [- <низ>]}.

Ако параметърът получава стойност след отговор на въпрос, ESTA ще покаже диалогова кутия с въпроса и падащ списък с имена на възможни опции или обяснителни текстове. Изборът може да стане от падащия списък или чрез избор на полета от илюстриращата картина. За целта, трябва предварително полетата да бъдат дефинирани и свързани към опциите чрез hotspot редактора, който е достъпен през командното меню Pictures Database.

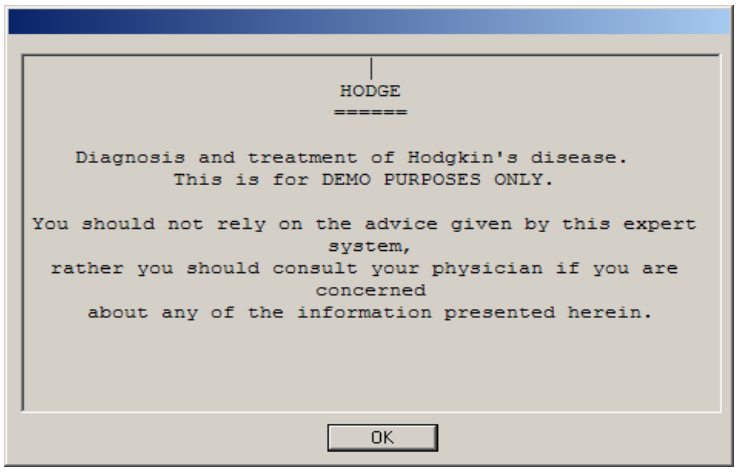
Пример. A5: от **cinema.kb**
parameter cinema : 'името на киното, в което показват вашия филм'
type category
options
odeon,
palace.
question 'В кое кино се показва вашия филм?'
На фиг. 7 е показан диалогът за параметъра cinema.



Фиг. 7. Пример за параметър тип category

1.3. Заглавия

Заглавията служат за представяне на всяка база знания, след зареждането и в ESTA. Дават възможност да се добавя и графична картина (picture) за презентирането им или илюстриране на диалога със системата по време на консултацията. На фиг. 8 е показан пример за заглавието на системата за диагностика на болестта на Хочкинс.



Фиг. 8. Пример за заглавие

1.4. Илюстрации (картини) в ESTA

ESTA поддържа илюстрации (картини) от следните типове: Windows metafiles (WMF), device dependent bitmaps (DDB) и device independent bitmaps (DIB).

Картините се съхраняват в картинна база данни (pictures database) наречена PICTURES.DBA. Имената на картините се използват като справка и по такъв начин се включват като част от ЕС. В базата данни могат да се използват общите функции като like Add (добавяне), Edit (редактиране) и Delete (изтриване).

Може да се добавят картини или чрез вмъкване (importing) от файл, или чрез копиране от Clipboard-а. Всички DDB, DIB и WMF по-големи от 64 KB се запаметяват във файл само за четене. Използвайки вертикалния и хоризонталния скролбар може да се избира мястото за картината в прозореца. Може да се настройва размера на картината или прозореца. ESTA включва също hotspot редактор, който позволява да дефинират избираеми полета от картината. Това дава възможност да се свързват с параметъра от тип категория и опциите им да се избират чрез тези полета от картината. Картините в ESTA могат да се използват като:

1. Стартова картина на ESTA системата;
2. Заглавие за базата знания;
3. Част от консултиращия диалог, за свързване с параметрите;
4. Част от съвета;
5. Част от списъка с действия в секцията showpic (picture).

Базата знания CAR.KB съдържа примери за всички споменати приложения на картините.

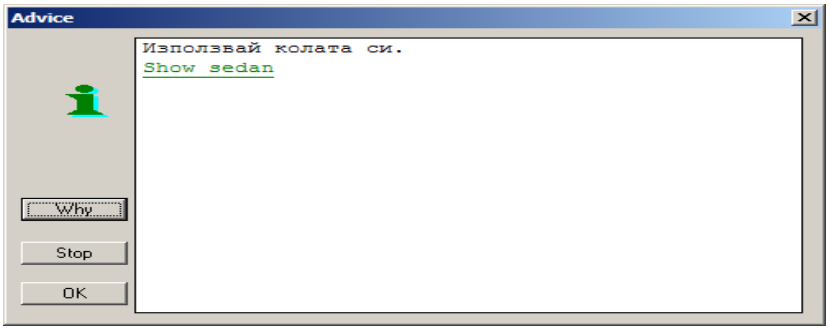
2. Действия

2.1. Съвет

Действието даване на съвет (advice) се изпълнява първо за променливите на всички параметри включени в текстовите изрази и се изписва резултатния текст в прозорец върху екрана. Могат да се включват и картинки в съвета. В такъв случай ще се появи хипервръзката към картинката със зададеното име в текста на съвета. Чрез кликване с мишката, тя може да се визуализира. Синтаксисът е:

<съвет> ::= **advice** <спецификация>
< спецификация > ::= <пункт_от_извода> {< спецификация >}
< пункт_от_извода > ::= текстов-израз | **picture** <име_на_картината>

Пример. А6: съвет даван от системата.
advice
‘Използвай колата си.’
picture 'sedan'



Фиг. 9. Съвет

2.2. Присвояване

Действието присвояване се използва за задаване стойности на параметрите. Синтаксисът е:
<присвояване> ::= **assign** <име-на-параметър> := <израз>

Би било добре формата на <израз> и резултатната стойност да съответстват на типа на параметъра. Може да се отдели секция, в която да извършва някаква начална инициализация на променливите, ако е необходимо.

Пример. 7

```
section start 'много къса секция за илюстриране на присвояването'
        assign n := 7 + 8 / 2
        advice 'Стойността на 7 + 8 / 2 е  ' n
        assign n := (7 + 8) / 2
        advice 'Стойността на  (7 + 8) / 2  е ' n
        'като / има по-висок приоритет отколкото +.' &
        'Бележка , че този параметър наистина променя стойността'
parameter n 'n'
type number
```

2.3. Извикване

Това действие се използва за извикване на вградени, или дефинирани от потребителя процедури, в ESTA. Аргументите на процедурата могат да бъдат изрази, които да кореспондират с техния тип.

Синтаксисът е:

```
<извикване> ::= call clear_all() |
        call clear_value(<име-на-параметър>) |
        call display(<име-на-файл>) |
        call hyperadvice(<име-на-файл >,<име>) |
        call restore_values(<име-на-файл >) |
        call save_values(<име-на-файл >) |
        call showpic(<име-на-картина>) |
        call sound(<Продължителност>,<Честота>) |
        call system(<низ>)
```

2.4. Поредица

Действието поредица на ESTA е предназначено да се консултира с нова БЗ. Ако е възможно, БЗ се разделя на няколко по-малки бази знания и така се повишава нагледността им. Синтаксисът е:
<верига> ::= **chain** <име-на-файл>

Стойностите на параметрите могат да се прехвърлят между базите знания чрез извикване на save_values (съхраняване на стойностите) и извикване на restore_values, преди първото действие в новата БЗ. Само параметри-променливи могат да бъдат прехвърлени между базите знания. Дефинициите се загубват, когато БЗ се свързват в поредицата.

Ако параметърът име-на-файл е празен текст, например *chain ‘ ‘*, ESTA ще се свърже към текущата БЗ с ефект на рестартирането и.

Пример. 8
Действието chain '**car.kb**' стартира консултация от БЗ **car.kb**. Тази нова БЗ ще бъде заредена в паметта и '**старата**' база знания по такъв начин ще бъде загубена.

2.5. Действие

Действието **do** е най-елементарния начин за прехвърляне на контрола към новата секция. Така се дава възможност за 'управлявано от данните' търсене пренесено между секциите. В литературата по ИИ това се нарича като 'права верига'. Синтаксисът е:

<действие> ::= **do** <име-на-секция>

Пример. 9
```
section start 'първата секция, която ще се изпълни'
  if answer = 'да' (do positive_section, do next_section)
  if answer = 'no' do negative_section
```

В зависимост от стойността на параметъра answer-отговор контролът се прехвърля transferred към секция positive_section или следващата я next_section, ако отговорът е ‘да’; или към секцията negative_section, ако отговорът е ‘не’.

2.6. Изпълнение на една избраните секции

Действието do_section_of взема един от аргументите на параметъра от тип category. Това е алтернатива на често използваните изрази като:
if < име-на-параметър-тип-категория > = <опция1> do < опция1>
if <име-на-параметър-тип-категория> = < опция2> do < опция2>
...

Вместо да се изписват всички if-оператори свързани към възможните опции, може просто да се напише:

do_section_of < име-на-параметър-тип-категория >

По време на консултацията стойността на параметъра ще определи на коя от секциите ще бъде прехвърлен контролът. Ако параметърът не е получил стойност, ESTA първо го изчислява и получената стойност се използва като указвате към съответната секция. Действията на do_section_of и на do действията е точно една и съща. Разликата е, че не е необходимо да се изписват всички еднакви изрази с do-оператори, а вместо това само с един израз. Трябва да се запомни, че имената на опциите в параметъра тип category трябва да съответства на имената на секциите в БЗ.

Пример. 10
Този пример е от базата знания CAR.KB:
```
section start главна секция'
do_section_of problem
 parameter problem : 'проблем с колата ви'
type category
explanation
'Намиране на проблема с вашата кола, е описан в тази база знания.'
options
starting_problems – проблем с пускането'
overheating - 'прегряване на двигателя'
smell_of_gasoline - 'колата мирише на бензин'
bad_running - 'двигателят се пуска лошо'
brakes - 'спирачките'
vibration – ‘прекалени вибрации'

wiper_motor - 'четките на мотора за чистачките’
light_problem - 'светлините’
```

horn_problem - 'неизправен клаксон'.
question
'Какъв е проблемът с вашата кола ' car ' ?'
picture 'car'

2.7. Изход

Действието **exit е** елементарен начин, за да се завърши консултацията за текущата БЗ. Синтаксисът е:
<exit> ::= **exit**

2.8. Спиране

Действието спиране - stop може да бъде използвано за оптимизиране на описаните в секцията правила. Изпълнението на действието спиране показва, че няма повече действия за изпълнение или булев израз е изчислен в съдържанието на секцията. Това действие помага на инженера по знанията да отхвърли повторението на объркани условия. Синтаксисът е:
<stop> ::= stop
Пример. 11. Елементарен пример, използващ спиране.
section start 'секция без действие'
stop
advice 'Този съвет никога не ще бъде даден докато не се изпълни действието stop '

3. Възможности на ESTA за връзка с други приложения на MS Windows.

ESTA поддържа DDE (Dynamic Data Exchange) интерфейс с други приложения като: електронни таблици (spreadsheets), бази данни (databases), текстови редактори (word processors). Възможно е да се изпращат или получават данни от приложения, или да се изпълняват команди. DDE интерфейса на ESTA прилича на макроезика на Word и Excel. Може да установява една или повече връзки с DDE сървъри.

DDE интерфейса поддържа следните процедури и функции:

- **dde_initiate**

Функцията: CH := dde_initiate(App,Topic) инициира връзка с приложението App. Приложението се идентифицира по името на неговия exe файл без разширение. App е низ. Темата topic е специфичен низ за приложението, което идентифицира нещо за връзката. Функцията връща номера на канала CH, по който става връзката и се използва от другите DDE функции и процедури.

Пример

Функцията, CH := dde_initiate('excel','test.xls'), инициира връзка с приложението MS EXCEL с таблицата 'test.xls' и връща номера на канала в променливата CH.

- **dde_execute**

Процедурата dde_execute(number,string) изпраща командата string по канала с номера на партньора за връзката. Синтаксиса на командата зависи от партньора.

Пример.

За извикване на Excel процедура: dde_execute(Ch,'[OPEN(\034test.xls\034]'), ще изпрати командата до Excel, в резултат, на което листа с таблицата test.xls ще бъде отворена. Кавичките се означават с тяхната ASCII стойност 034.

- **dde_poke**

Процедурата dde_poke(CH,Item,Data) изпраща данните Data като низ към Item, който също е низ по канал с номер CH към партньора.

Пример

За извикване на Excel: dde_poke(CH,'R2C4','hello'), ще вмъкне низа hello в клетката R2C4.

- **dde_poke_number**

Процедурата dde_poke_number(CH,Item,Data) изпраща данните Data определени като число, към Item, определен като стринг, по канал с номер CH към партньора за връзката.

Пример

За извикване на Excel: dde_poke_number(CH,'R2C4',123) ще вмъкне числото 123 в клетката R2C4.

- **dde_request**

Функцията STR := dde_request(Ch,Item) връща низа STR относно низа Item към партньора за връзката, идентифициран чрез канал с номер Ch.

Пример

Ако партньора е Excel, то Item може да бъде клетка в листа с таблицата:
assign X := dde_request(CH,'R2C5')
Стойността на клетката R2C5 се връща като низ и се присвоява на X, който е текстов параметър**.**

- **dde_request_number**

Функцията Number := dde_request_number(Ch,Item) връща числото Number относно низа Item към партньора за връзката, идентифициран чрез канал с номер Ch.

Пример

Ако партньора е Excel, Item може да бъде клетка в листа с таблицата:
assign X := 10 + dde_request_number(CH,'R2C5')
Стойността на клетката R2C5 се връща като число и се присвова на X, който е текстов параметър**.**

- **dde_terminate**

Процедурата dde_terminate(CH) ще прекъсне връзката, идентифицирана чрез номера на канала CH.

- **dde_terminate_all**

Процедурата dde_terminate_all прекъсва всички DDE връзки.

ЗАКЛЮЧЕНИЕ

ESTA е удобна среда за разработване на експертни системи, която дава възможност за приложение в обучението на студентите в университетите.

Това се дължи на богатия набор от вградени средства, позволяващи лесно извличане, редактиране, формализиране на знанията за предметната област и избор на механизъм за извод с подходяща стратегия за конкретния случай.

От две години тази шел-среда се използва в обучението на студентите от Шуменския университет по дисциплината „Експертни ситеми” при разработването на курсови и дипломни проекти. Това повиши ефективността при провеждане на практическите занятия и позволи да се подобрят резултатите от тях.

ЛИТЕРАТУРА

1. **Ненков** Н.В, Експертни системи, Университетско издателство, Шумен, 2006.
2. www.visual-prolog.com.

ВЪРХУ ЕДНА ОПТИМИЗАЦИЯ НА ТЪРСЕНЕТО В ХЕШ-ТАБЛИЦА С ОТВОРЕНО АДРЕСИРАНЕ

ВАЛЕНТИНА СП. ДЯНКОВА, РОСИЦА П. ХРИСТОВА

ON OPTIMIZING THE SEARCH IN THE OPEN-ADRRESING HASH TABLE

VALENTINA SP. DYANKOVA, ROSSICA P. CHRISTOVA

The article talks about the realization of a hash table, which is a table with open addressing in the sequential section of memory. It gives the opportunity for reducing the number of linear probes when a particular data element is not inserted in the hash table.

KEY WORDS: open-addressing hash table, collision, element search,remove an element.

1. ВЪВЕДЕНИЕ.

При писането на програми, моделиращи по-сложни обекти, процеси и явления, се конструират по-сложни структури от данни. Това налага задаване на начина на представяне на информацията (нейната структура), определяне на връзките между отделните й елементи, както и дефиниране на необходимите операции за нейната обработка. Правилният избор на подходяща структура е необходимо условие за постигане на оптимално решение. Естествено възниква въпросът за реализация използваната структура от данни. При реализирането на една и съща структура могат да бъдат използвани различни подходи, адаптирани към конкретните изисквания за бързодействие, използвана памет и др.

В настоящата статия се разглежда една реализация на хеш-таблица с отворено адресиране в последователно разпределена памет, даваща възможност за намаляване на броя на линейните проби при установяване на факта, че даден елемент не е включен в хеш-таблицата .

2. ПОНЯТИЯ СВЪРЗАНИ С ХЕШ-ТАБЛИЦИ С ОТВОРЕНО АДРЕСИРАНЕ.

Таблицата е съвкупност от еднотипни елементи, всеки от които е изграден от ключ (идентификационна част) и тяло (информационна част). Стойността на ключа еднозначно идентифицира конкретен елемент. Стойностите на информационната част не са съществени от гледна точка на организацията и обработката на таблицата, поради което в предложената реализация не е зададен конкретен тип за нея.

В хеш-таблицата достъпът до елемент се осъществява чрез трансформиране на ключа на елемента в неговия адрес, т. е. търсенето може да се определи като изображение $\text{hash: K} \rightarrow \text{A}$, което се нарича хеш-функция. Тъй като за представянето на хеш-таблица в последователно-разпределена памет се използва масив, трансформацията на ключ в адрес се свежда до трансформиране на ключ в индекс на масива.

Изборът на добра хеш-функция е гаранция за равномерно разпределение на елементите в хеш-таблицата, но тъй като това няма отношение към целта на настоящата статия, то считайки без ограничение на общността, че ключовете на елементите са естествени числа, ще бъде използвана класическата хеш-функция

$hash(k) = k\%N$, където k е ключ на елемента, а N - броя на елементите в масива.

Тъй като мощността на множеството K е по-голяма от възможния брой на елементите в масива, то е възможно да възникне ситуация, в която два елемента с различни ключове $k_1 \neq k_2$ да претендират за едно и също място в масива, т. е. $hash(k_1) = hash(k_2)$. Такива елементи се наричат синоними, а самото явление – колизия. Проблемът за разрешаване на колизиите има различни решения, но в статията ще бъде разгледан методът на линейното отворено адресиране. При този

метод, ако елемент претендира за място, заето от друг елемент, масивът се преглежда последователно за ново свободно място. Последователното търсене може да се реализира с функцията: $rehash(i) = (i+1)\%N$

Основните операции в хеш-таблици са:

Търсене на елемент с ключ k - с помощта на хеш-функцията $hash(k) = a_1$ се пресмята индекса a_1 на позицията в масива, където би трябвало да се намира този елемент; но ако в нея има елемент с различен ключ, проверката се извършва за позиция $a_2 = rehash(a_1)$. Аналогични действия се извършват и с позиция a_2. Този процес на линейни проби продължава докато: бъде намерен елемент с ключ k (успешен край), бъде достигнат елемент с нулев ключ (неуспешен край) или чрез обхождане на цялата таблица бъде установено, че елемент с такъв ключ отсъства.

Добавяне на елемент с ключ k - ако се установи, че елемент с такъв ключ отсъства от таблицата и в нея има свободно място, то той се добавя в таблицата.

Изтриване на елемент с ключ k - след успешен край на операцията търсене на елемент с ключ k и установяване на неговото местоположение в позиция i, се нулира ключа на елемента в тази позиция (по-долу се цитира, като очевидния метод за изтриване на елемент).

3. Проблеми за разрешаване.

Още Кнут във фундаменталния си труд [1] пише: “Многие програмисты свято верит в алгоритмы и очень удивляются, обнаружив, очевидный способ удаления записей рассеянной таблицы не работает.”.

Прегледа на класическата литература по въпроса показва, че операцията изтриване от хеш-таблици с отворено адресиране или не се коментира [2], [3], [4], или се коментира по някой от следните начини:

- Маркирането на елемента с очевидния метод скъсва веригата на синонимите[5], [8];
- Въвежда се спецификатор с три състояния: елемента е зает; елемента е свободен, като никога не е бил заеман; елемента е свободен, но преди това е бил зает [6], [7].

Никой от горепосочените автори не предлага реализация на изтриване на елемент, единствената предлагана реализация е по очевидния метод в [8].

Ефективността при търсене на предложения метод с три състояния [6], [7] се коментира като незадоволителна. В последната за периодично подобряване на скоростта на търсене се предлага процесът “събирач на боклук”.

Освен проблема разгледан по-горе, така дадените дефиниции на операциите изтриване и търсене пораждат следните въпроси:

2.1. Как се отразява изтриването на елемент от веригата линейни проби върху търсенето на елемент, влизащ в състава на тази верига?

2.2. Достатъчно основание ли е достигането на нулев ключ във веригата линейни проби при търсене на елемент, за да се твърди, че търсеният елемент не е в таблицата?

2.3. Необходимо ли е обхождане на таблицата до достигане на празен елемент (който никога не е запълван), за да бъде установено, че търсеният елемент отсъства?

4. Един пример

Като илюстрация на повдигнатите въпроси може да бъде разгледан следния пример: в хеш-таблица с размер 13 се разполагат елементи с ключове:

- 14 (hash(14)=14%13=1);
- 16 (hash(16)=16%13=3);
- 29 (hash(29)=29%13=3, rehash(3)=4%13=4);
- 55 (hash(55)=55%13=3, rehash(3)=4%13=4, rehash(4)=5%13=5);
- 21 (hash(21)=21%13=8);
- 35 (hash(35)=35%13=9);
- 49 (hash(49)=49%13=10);
- 50 (hash(50)=50%13=11).

0	1	2	3	4	5	6	7	8	9	10	11	12
0	**14**	**0**	**16**	**29**	**55**	**0**	**0**	**21**	**35**	**49**	**50**	**0**

При прилагане на разпространения в литературата алгоритъм за търсене за елементи с ключове 14, 20, 55, 42, 48 се получават следните резултати:

- за 14: чрез прилагане на хеш-функцията се получава индекс 1, където се намира търсения ключ, т. е. търсенето завършва успешно;
- за 20: чрез прилагане на хеш-функцията се получава индекс 7, където стои ключ 0, т. е. мястото е празно и търсенето завършва неуспешно;
- за 55: чрез прилагане на хеш-функцията се получава индекс 3, където стои ключ 16≠55. Следвайки класическия алгоритъм след прилагане на метода на линейните проби 2 пъти, се достига до ключа 55 и търсенето завършва успешно;
- за 48: чрез прилагане на хеш-функцията се получава индекс 9, където стои ключ 35≠48. Прилагайки метода на линейните проби3 пъти, се достига до ключ 0, т. е. до празно място и търсенето завършва неуспешно. В този случай може да бъде поставен следния въпрос: не може ли още при получаването на индекс 9 да направим извода за неуспешен изход на търсенето, тъй като от процеса на разполагане на елементите в таблицата се забелязва, че в позиция 9 не е възниквала колизия и няма как да се очаква търсения ключ да се появи във веригата линейни проби за тази позиция (въпрос 2.3).

За илюстрация на повдигнатите въпроси 2.1. и 2.2. може да се разгледа търсенето на елемент с ключ 55 след изтриването на елемент с ключ 29. Разположението на елементите след операцията изтриване е следното:

0	1	2	3	4	5	6	7	8	9	10	11	12
0	**14**	**0**	**16**	**0**	**55**	**0**	**0**	**21**	**35**	**49**	**50**	**0**

Тогава при прилагане на алгоритъма за търсене на елемент с ключ 55 и обхождайки елементите се достига до индекс 4, където стои ключ 0, т. е. мястото е празно и съгласно класическия алгоритъм (търсене до откриване на ключа или празно място) търсенето би трябвало да завърши неуспешно, което противоречи на действителното разположение на елементите в таблицата.

5. РЕШЕНИЕ НА ПРОБЛЕМА

Едно коментирано вече разрешение на този проблем е въвеждането на спецификатор с три състояния (реализация е публикувана в [9]). При него операцията изтриване се реализира като за ключ се записва стойност, различна от нулевия ключ и непринадлежаща на множеството от допустими стойности за ключа (напр. -1). Тогава, ако в процеса на търсене на елемент се попадне върху ключ, с който елемента е маркиран за изтрит, се продължава с метода на линейните проби. Това осигурява успешен край на търсенето, ако преди това от веригата линейни проби е бил изтрит елемент, но допуска извършването на линейни проби, дори и когато това не е необходимо.

Основния проблем, на който се предлага решение в статията е реализацията на алгоритъм за търсене на елемент в хеш-таблица с отворено адресиране, при който се намалява броя на линейните проби при установяване на факта, че даден елемент не е включен в хеш-таблицата, а изключването на елемент няма неприятни последствия върху алгоритмите за търсене и включване на елемент във вътрешните вериги на синонимите при разрешаване на колизии.

Решението на този проблем е за всяка позиция да се пази информация дали на съответното място е възниквала колизия при добавянето на елементи. Това води до идеята за алтернативен масив ph от булеви стойности:

$$ph[i] = \begin{cases} true, & ако\ \ на\ \ i-то\ \ място\ \ е\ \ имало\ \ колизия \\ false & в\ \ противен\ \ случай \end{cases}$$

Тогава разполагането на разгледаните по-горе елементи ще предизвиква следните промени в двата масива:

- 14 (hash(14)=14%13=1):

	0	1	2	3	4	5	6	7	8	9	10	11	12
T	**0**	**14**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

ph	false	false	false	false	false	false	false	false	false	false	false	false	false

- 16 (hash(16)=16%13=3):

	0	1	2	3	4	5	6	7	8	9	10	11	12
t	**0**	**14**	**0**	**16**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
ph	false	false	false	false	false	false	false	false	false	false	false	false	false

- 29 (hash(29)=29%13=3; rehash(3)=4%13=4);

	0	1	2	3	4	5	6	7	8	9	10	11	12
t	**0**	**14**	**0**	**16**	**29**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
ph	false	false	false	true	false	false	false	false	false	false	false	false	false

- 55 (hash(55)=55%13=3, rehash(3)=4%13=4, rehash(4)=5%13=5):

	0	1	2	3	4	5	6	7	8	9	10	11	12
t	**0**	**14**	**0**	**16**	**29**	**55**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
ph	false	false	false	true	true	false	false	false	false	false	false	false	false

- 21 (hash(21)=21%13=8):

	0	1	2	3	4	5	6	7	8	9	10	11	12
t	**0**	**14**	**0**	**16**	**29**	**55**	**0**	**0**	**21**	**0**	**0**	**0**	**0**
ph	false	false	false	true	true	false	false	false	false	false	false	false	false

- 35 (hash(35)=35%13=9):

	0	1	2	3	4	5	6	7	8	9	10	11	12
t	**0**	**14**	**0**	**16**	**29**	**55**	**0**	**0**	**21**	**35**	**0**	**0**	**0**
ph	false	false	false	true	true	false	false	false	false	false	false	false	false

- 49 (hash(49)=49%13=10):

	0	1	2	3	4	5	6	7	8	9	10	11	12
t	**0**	**14**	**0**	**16**	**29**	**55**	**0**	**0**	**21**	**35**	**49**	**0**	**0**
ph	false	false	false	true	true	false	false	false	false	false	false	false	false

- 50 (hash(50)=50%13=11):

	0	1	2	3	4	5	6	7	8	9	10	11	12
t	**0**	**14**	**0**	**16**	**29**	**55**	**0**	**0**	**21**	**35**	**49**	**50**	**0**
ph	false	false	false	true	true	false	false	false	false	false	false	false	false

При това състояние на масивите търсенето на елемент с ключ 48 ще доведе до прилагане на хеш-функцията над ключа 48 – hash(48)=48%13=9. Тъй като на това място не стои елемент с търсения ключ и $ph[i] = false$, то не е необходимо прилагането на линейните проби и може да бъде направен извод за неуспешен край на търсенето.

6. РЕАЛИЗАЦИЯ

Една реализация на Pascal на предложеното решение:

```
const TabSize=13;
    NilKey=0;
type  KeyType=integer;
    InfoType=<име на тип>|<дефиниция на тип>
    ElementType=record
            key: KeyType;
            info: InfoType
            end;
    Index=0..TabSize;
    Hash =function (K:KeyType):Index;
    ReHash= function (I:Index):Index;
```

```
    HashTable=record
            T:array [Index] of ElementType;
            Ph:array [Index] of Boolean;
            Free:Index; {указва брой свободни места}
            H:Hash; {хеш-функцията}
            R:ReHash; {рехеширащата функция}
            end;
function Hashf (K:KeyType):Index; far;
begin  Hashf:= K mod TabSize  end;
function Rehashf (I:Index):Index; far;
begin  Rehashf:= (I+1) mod TabSize  end;
procedure Init ( var HT:HashTable); {инициализация на хеш-таблицата}
var i: Index;
begin  with HT do  begin
     for i:=0 to TabSize–1 do
        begin  T[i].Key:=NilKey;   Ph[i]:=false;  end;
     Free:=TabSize;  H:= Hashf;  R:= Rehashf;
end;            end;
function Is_Full (HT:HashTable):boolean; {проверка за пълна таблица}
begin  Is_Full:= HT.Free=0  end;
function Search (K:KeyType; HT:HashTable):integer;
{търсене на елемент в хеш-таблица – връща номера на индекса  или –1 }
var i, j:Index;  b:boolean;
begin  b:=false;  i:=HT.H(K);  j:=i;
  with HT do  begin
     while (T[i].Key<>K) and Ph[i] and not b do
        begin  i:=R(i);  b:= i=j  end;
     if T[i].Key=K then Search:=I  else Search:=-1;
end;         end;
procedure Del (K:KeyType;var HT:HashTable); {изтрива елемент по ключ}
var i:integer;
begin  i:= Search(K, HT);
  if  i>=0 then  {елемента е открит на позиция с индекс i}
     begin  HT.T.Key:=NilKey;  HT.Free:=HT.Free+1  end;
end
procedure Insert (E:ElementType;var HT:HashTable); {добавя елемент}
var i:integer;
begin  i:= Search(E.Key, HT);
  if  (i<0) and not Is_Full(HT) then
    with HT do  begin  i:=H(E.Key);
      while T[i].Key<>NilKey do
         begin  Ph[i]:=true;  i:=R  end;
      T[i]:=E;  Free:=Free–1;
end;            end;
```

ЗАКЛЮЧЕНИЕ

Предлаганата реализация на хеш-таблица с отворено адресиране в последователно разпределена памет може да се използва с прилагането на произволни хеш-функции и произволен начин за обработване на колизиите, като във всеки от тези случаи съществено се намалява броят на пробите при възникването на колизии, особено ако търсения елемент не е от хеш-таблицата.

ЛИТЕРАТУРА

1. **Кнут** Д. Искусство програмирования для ЭВМ. т.3. Сортировка и поиск. Изд. Мир, М., 1978.
2. **Мейер** Б., К.Бодуэн. Методы программирования. т.2. Изд. Мир, Москва, 1982.
3. **Рейнголд** Э., Ю.Нивергельт, Н.Део. Комбинаторные алгоритмы. Теория и практика. Изд. Мир, М., 1980.
4. **Амерал** Л. Алгоритми и структури от данни в C++. ИК СОФТЕХ, 2001.
5. **Наков** П., П.Добриков. Програмиране=++Алгоритми. Изд. TopTeam Co., С., 2002.
6. **Шишков** Д. и др. Структури от данни. Изд. Интеграл, Добрич, 1995.
7. **Смит** Т. Принципи и методи на програмирането с PASCAL. Изд. Техника, С., 2001
8. **Азълов** П. Програмиране. Основен курс. Изд. АСИО, С., 1995.

WEB БАЗИРАНА ИНФОРМАЦИОННА СИСТЕМА „СТУДЕНТ”

Антон Г. Димитров, Петър А. Милев, Найден В. Ненков

WEB BASED INFORMATION SYSTEM “STUDENT”

Anton G. Dimitrov, Petar A. Milev, Nayden V. Nenkov

The paper presents a web based information system “Student”. The developed software codes enable the management of the activity of student inspector in local networks and Internet.

KEY WORDS: web based information system

УВОД

Традиционните начини за обработка на документацията свързана със студентите във висшите училища извършвана от инспекторите по студентските въпроси създава множество неудобства, които отнемат ценно работно време на служителите. Автоматизирането на този процес е свързано с анализ на тяхната дейност и документооборот и последващо моделиране и нормализиране на релационна база от данни (РБД) „Студент” [4]. Изграждането на web базирана информационна система (ИС), виртуален аналог на реална дейност, например извършено в [5], може да предостави възможност за on-line управление и координация на студентските информационни единици [3]. На тази основа целта на настоящата статия е разработването на web базирана ИС „Студент”. За постигането й първо е разгледан универсалния комуникационен модел клиент-сървър основан на web технологията [2], след това са представени програмно реализираните основни моменти в работния процес на инспектора по студентските въпроси, реализирани с помощта на ИС „Студент”. В заключение са направени изводи за получените резултати.

УНИВЕРСАЛЕН КОМУНИКАЦИОНЕН МОДЕЛ КЛИЕНТ-СЪРВЪР, ОСНОВАН НА WEB ТЕХНОЛОГИЯТА

Структурата на универсалния комуникационен модел клиент-сървър е показана на Фиг. 1. Тя е основана на web технологията и включва две основни части. Първата обхваща всичко онова, което се намира между клиента и HTTP сървъра. Тази част е стандартизирана, независима е от платформите, основава се на Интернет услугите, и е в състояние да поддържа мрежи с нисък дебит. Втората част се отнася до всичко онова, което се намира след HTTP сървъра. В тази част се намират елементи и се прилагат технологии, специфични за класическата схема на клиент-сървър модела [2]. Клиентът управлява потребителския интерфейс и контролира въведените данни, с което се предотвратява излишния трафик по мрежата. Web сървърът (HTTPD) свързва данните между клиента и сървъра за данни. В частта сървър за данни се извършват обработки над данните и заявките (обикновено SQL) и се осъществява връзка със сървъра на базата от данни (БД) за достъп до данните. От страна на сървърите обработките могат да бъдат разпределени между няколко физически машини, а не върху една единствена, с което се облекчава тяхното натоварване и се увеличават ресурсните им способности да обработват значително количество заявки.

Фиг. 1 Схема на универсален клиент-сървър комуникационен модел

Действията на едно приложение в средата на универсален клиент-сървър се състоят от следната последователност от операции:

1. Установява се връзка между HTTP сървъра и web клиента и се извлича съответната страница.
2. Данните се въвеждат в HTML формуляри, изобразени от web браузър.
3. Скриптов език контролира на място въведената информация (Javascript, VBScript).
4. Заявката се изпраща към HTTP сървър по CGI метод.
5. Обработват се данните от заявката в сървъра и се генерират SQL заявки.
6. Достъп до данните посредством сървъра на БД (SQL сървър).
7. Генерира се HTML страница с получените данни. Страниците се генерират от сървър, за да бъдат интерпретирани от клиент.
8. Изпращат се резултатите при клиента и се визуализират от браузър.

Универсалният клиент-сървър извежда на преден план способността на един клиент, независимо от неговата платформа да контактува конкурентно с произволна конфигурация на сървър за БД посредством web технология или HTTP сървър. Web браузърите предоставят лесен за усвояване интерфейс и относително прости средства за програмиране, достъпни за широк кръг от потребители. Промени в структурата на БД, или процедурите за обработка на данните при сървъра, или промяна в неговата конфигурация не оказват влияние на клиентската част. Използват се две основни техники за осъществяване на достъп до информацията на БД посредством web технологията. Първата се базира на генерирането на статични HTML страници. Втората се основава на създаването на динамични БД, достъпни непосредствено чрез web. Трети вариант е възможен при комбиниране на първите два.

Един от често използваните програмни езици за създаване на интерактивни и динамични web сайтове е PHP [1], [7]. PHP е скриптов език, чиито код се обработва в момента, в който сървърът изпълнява клиентската заявка. Той осигурява подръжката на различни БД, например като MySQL [1], [6]. Съчетанието на PHP с MySQL предоставя мощен инструмент за изграждане на web базирани ИС. С помощта на този подход е изградена и представената в настоящата статия информационната система „Студент”.

УПРАВЛЕНИЕ НА ОСНОВНИЯ НАБОР ОТ ДОКУМЕНТИ И ЗАЩИТА НА WEB БАЗИРАНАТА ИНФОРМАЦИОННА СИСТЕМА „СТУДЕНТ”

В съответствие с документооборота свързан със студентските информационни единици съществуват и различни алгоритми за тяхното управление.

УПРАВЛЕНИЕ НА РЕЛАЦИОННА БАЗА ОТ ДАННИ „СТУДЕНТ”

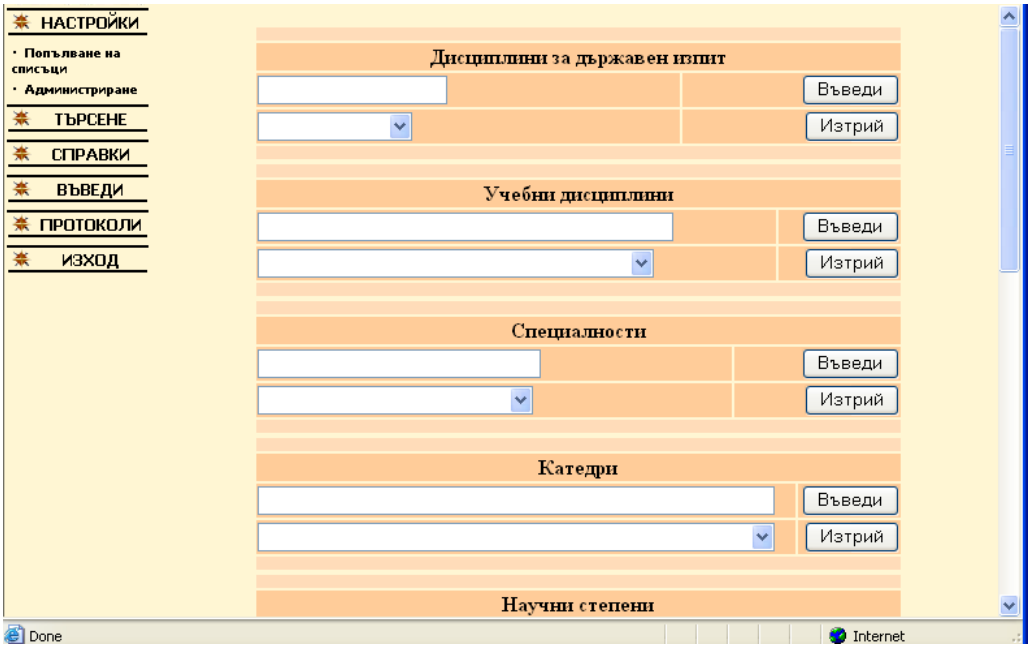
Принципът на използване на РБД “Студент” от една система за управление (СУ) на РБД, каквато се явява ИС “Студент”, е следният:

1. Клиент изпраща искане към ИС.
2. Системата от своя страна обработва данните и се обръща към MySQL сървъра:
   2.1. Проверява се за наличие на потребителски достъп.
   2.2. Избира се БД, която ще се използва.
   2.3. Сървърът изпълнява заявката и връща резултата на СУРБД.
   2.4. Затваря се връзката към MySQL.
3. ИС обработва получените данни.
4. Резултатите се изпращат на клиента.

ВЪВЕЖДАНЕ/ИЗТРИВАНЕ НА СЛУЖЕБНА ИНФОРМАЦИЯ

Преди да започне същинската работа на web базираната ИС „Студент”, а и по време на нейната обичайна работа, е необходимо въвеждането/изтриването на данни -специалности, форма на обучение, дисциплини, имена на преподаватели и др. – свързани със служебната информация. Това става по следния начин:

1. Инспектора по студентски въпроси влиза в системата със съответното си потребителско име и парола.
2. От менюто се избира НАСТРОЙКИ -> Попълване на списъци
3. В дясната част на прозореца се появяват основните списъчни полета за вмъкване и изтриване (Фиг.2).



Фиг. 2. Въвеждане/изтриване на служебна информация

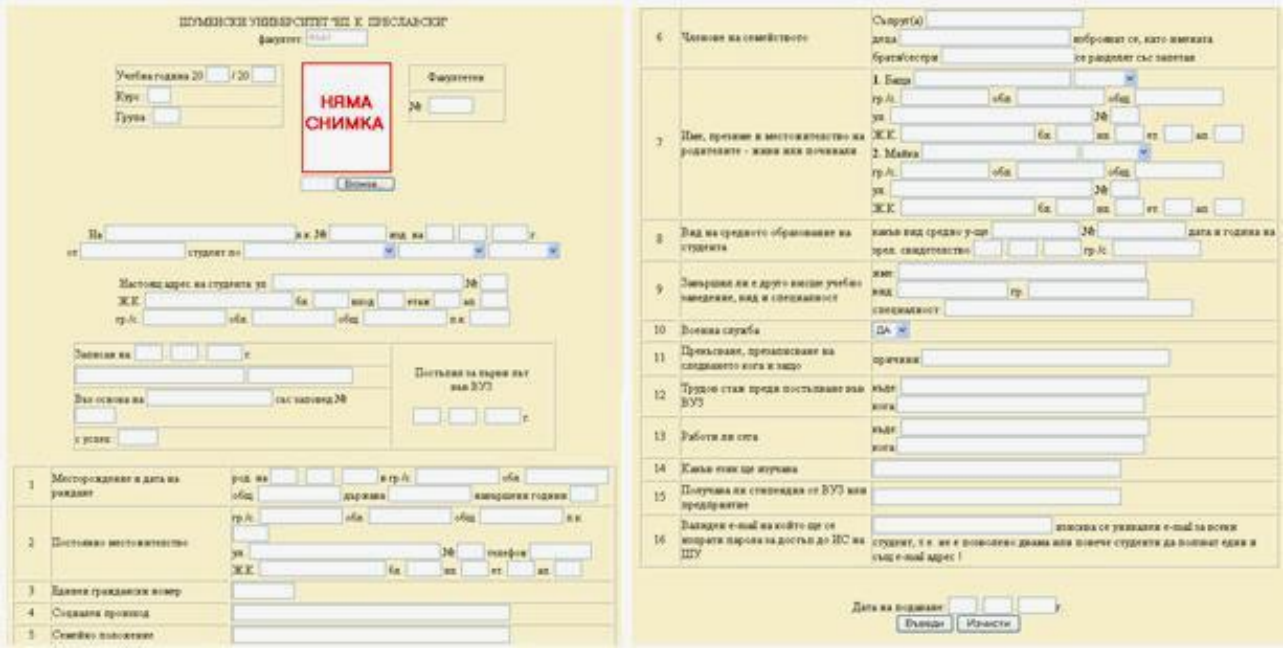
4. Въвеждането се извършва като се попълни съответното текстово поле и се натисне бутона “Въведи”.
5. Изтриването се извършва като се избере от падащия списък съответната стойност, която ще се изтрие и се натисне бутона “Изтрий” в съответната страница.

ВЪВЕЖДАНЕ/ТЪРСЕНЕ ДАННИ ЗА СТУДЕНТИТЕ

За въвеждане на студент в ИС е необходимо попълването на входния документ “Именник”. Разглеждат се два случая - дали студента се въвежда за първи път или вече е въведен в БД и се записва в по-горен курс.

Първо се разглеждат действията, които трябва да се извършат при записването на студент за първи път:

1. От менюто се избира ВЪВЕДИ\Именник за нов студент.
2. В дясната част на прозореца се появява формуляр, в който трябва да се попълнят данните за студента, (Фиг. 3).

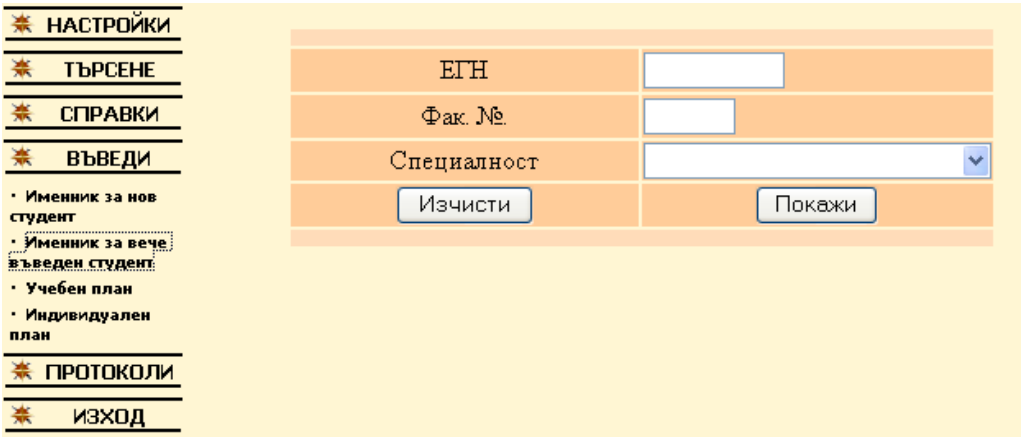


Фиг. 3. Документ „Именник”

3. След като се въведат данните във формуляра, се натиска бутон “Въведи”, намиращ се най-отдолу на формуляра.
4. Ако всички данни са попълнени коректно, студентът се записва в БД, но ако има непопълнени задължителни полета или некоректни такива, системата оцветява тези полета в червено. След това се повтаря стъпка 3.

При запис на студент в по-горен курс неговите данните вече са въведени в ИС. Изпълнява се следната последователност от действия:

1. От менюто се избира ВЪВЕДИ -> Именник за вече въведен студент.
2. В дясната част на прозореца се появяват полета за търсене по ЕГН, Фак. № и специалност, (Фиг. 4)



Фиг. 4. Търсене на въведен студент

3. След като се въведат данни в поне едно от полетата, се натиска бутона “Покажи”. Системата ще извърши търсене на студенти по въведените данни.

▪ Ако системата не намери студент, отговарящ на тези критерии, ще изведе съответно съобщение;

▪ Ако открие един студент, системата автоматично ще премине към стъпка 4;

▪ Ако намери двама или повече студенти, отговарящи на зададените критерии, ще се появи списък с резултатите. Достатъчно е да се избере името на търсения студент и системата ще премине към стъпка 4, (Фиг. 5).

ЕГН	Фак. No.	Специалност	Име	Курс	Научна степен	Учебна година
8112302880	410	Информатика	Антон Георгиев Димитров	4	бакалавър	2003/2004
8010078244	512	Информатика	Първо Второ Трето	1	бакалавър	2000/2001
8112302880	410	Информатика	Антон Георгиев Димитров	3	бакалавър	2002/2003
8112302880	410	Информатика	Антон Георгиев Димитров	2	бакалавър	2001/2002
8112302880	410	Информатика	Антон Георгиев Димитров	1	бакалавър	2000/2001

Фиг. 5. Търсене на въведен студент

4. В дясната част на прозореца се появява формуляр, в който данните са попълнени и потребителят трябва да ги актуализира с тези от документа, представен от студента.

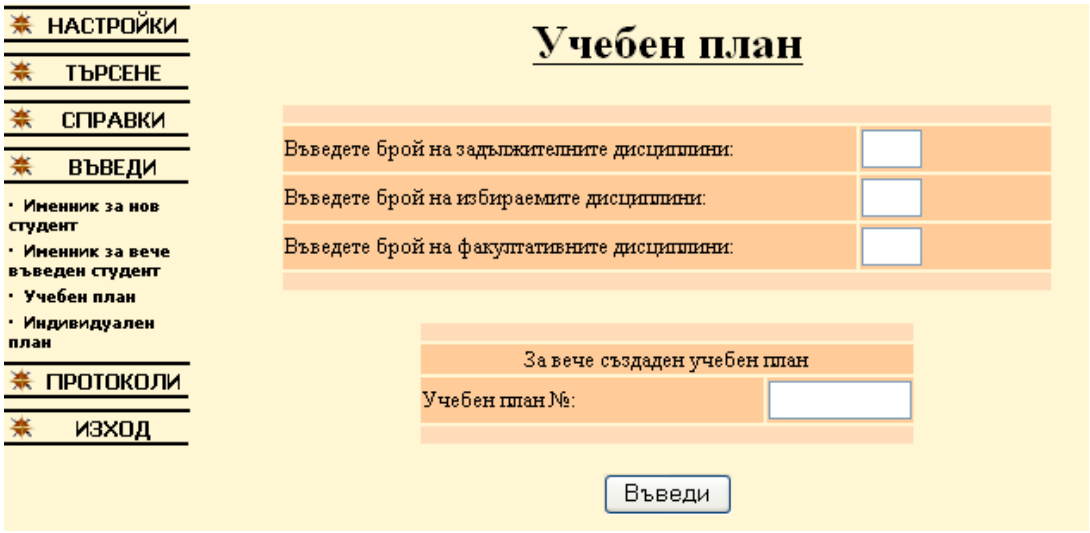
5. След актуализиране данните във формуляра, се натиска бутон “Въведи”, намиращ се най-отдолу на формуляра.

Ако всички данни са попълнени коректно, студента се записва в БД, но ако има непопълнени задължителни полета или некоректни такива, системата оцветява тези полета в червено. След това се повтаря стъпка 5.

РАЗРАБОТКА НА УЧЕБНИ ПЛАНОВЕ

За да се въведат учебните планове на студентите, трябва да се следват следните стъпки:

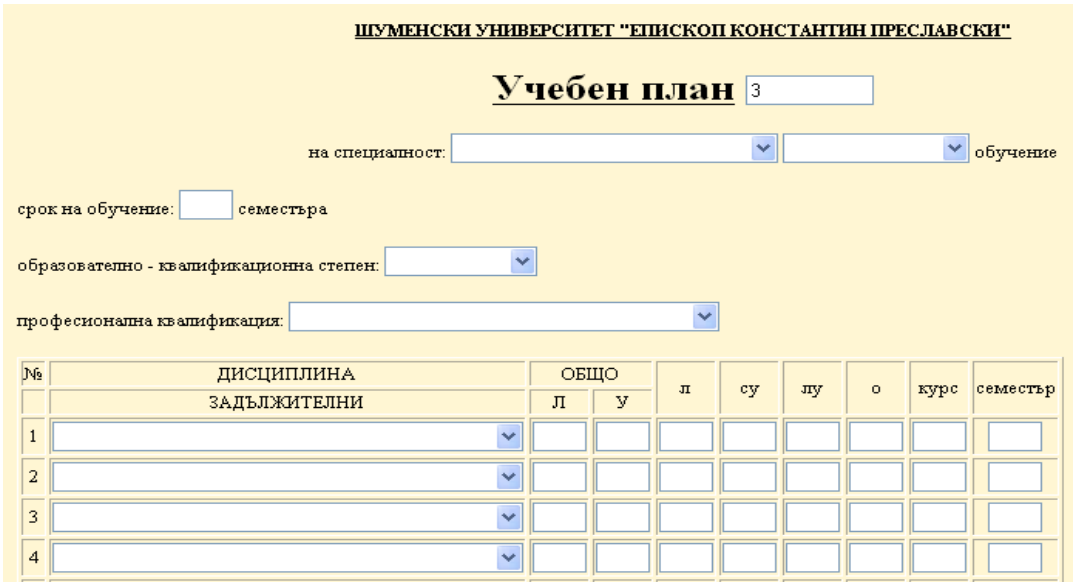
1. Въвеждането на нов или редактирането на вече съставен учебен план, става като се избере от менюто ВЪВЕДИ -> Учебен план.
2. В дясната страна на прозореца се визуализират четири полета за попълване, (Фиг. 6).



Фиг. 6. Разработка на учебен план

- Ако се създава нов план, то се попълват първите три полета и се натиска бутона „Въведи", който препраща на стъпка 3.
- Ако се редактира вече въведен учебен план, трябва да се запише номера на плана, който се актуализира. След като се натисне бутона „Въведи", ИС препраща към стъпка 3.

3. Преминаване във форма за създаване/актуализиране на учебен план, (Фиг. 7).



Фиг. 7. Създаване/актуализиране на учебен план

- Ако се въвежда нов учебен план, то формуляра ще бъде празен. След коректно попълване се натиска бутона "Въведи" и данните ще бъдат записани в БД. Ако не са

попълнени коректно, системата ще презареди страницата, като оцвети в червено полетата с некоректни данни.

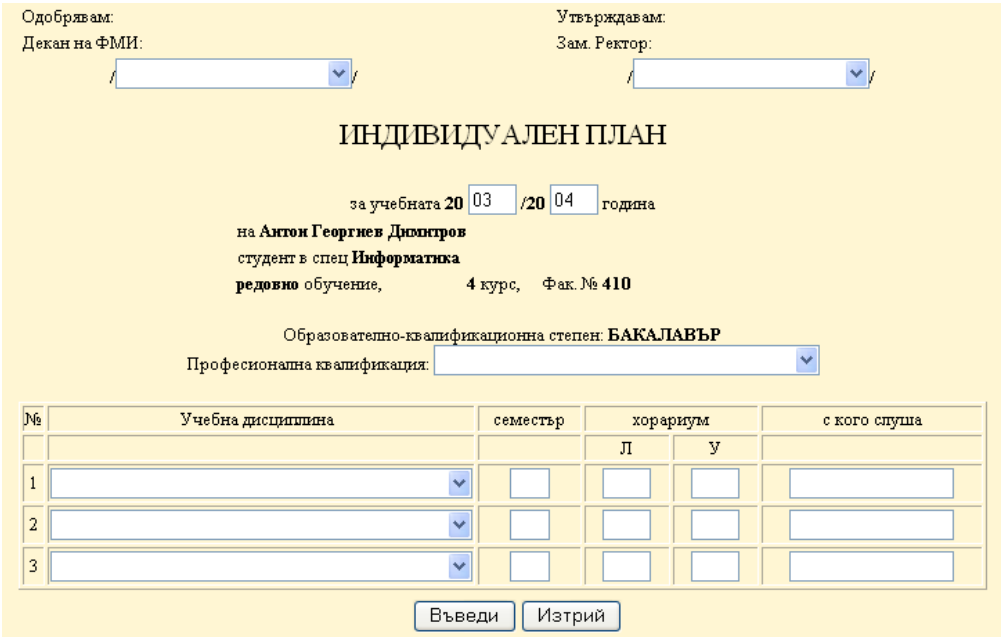
▪ Ако се запише номер на вече създаден план, който потребителят желае да редактира, формата която ще се изобрази ще бъде попълнена с данните от този план.

4. При редакция на информацията и натискане на бутона “Въведи”, данните в базата ще се актуализират. При натискане на бутона “Изтрий” – учебният план ще бъде унищожен.

РАЗРАБОТКА НА ИНДИВИДУАЛЕН ПЛАН

За да се създаде индивидуален план на студент е нужно да се следват следните стъпки:

1. От менюто ВЪВЕДИ се избира Индивидуален план.
2. В дясната част на прозореца се появява формуляр, който е празен, (Фиг. 8).



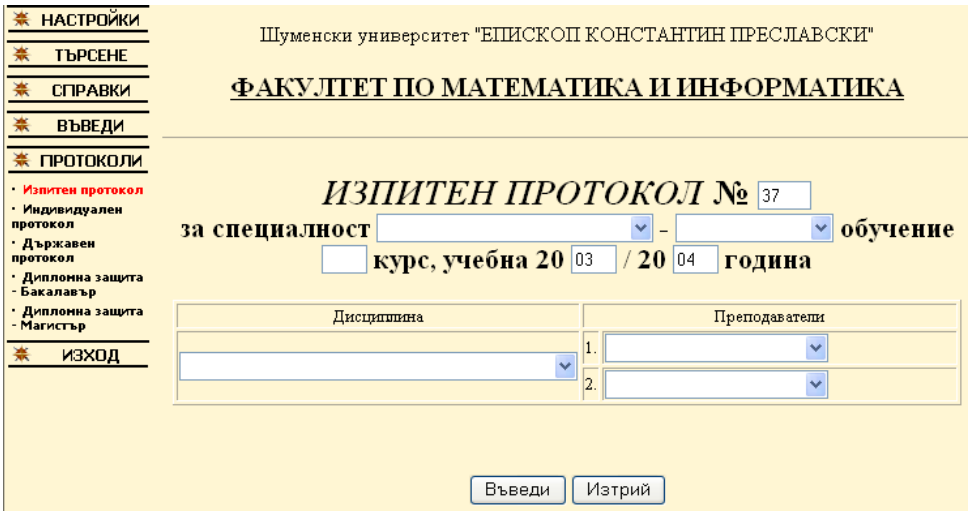
Фиг. 8. Формуляр за разработка на индивидуален план

3. След попълването му и натискане на бутона “Въведи”, индивидуалният учебен план вече е записан в базата данни. Но, ако е попълнено некоректно някое от полетата, тогава ИС връща в същият прозорец, като оцветява полетата с некоректни данни в червено.

РАЗРАБОТКА НА ИЗПИТЕН ПРОТОКОЛ

За да се създаде изпитен протокол за даден курс от студенти по определена специалност, трябва да се следват следните стъпки:

1. Избира се меню ПРОТОКОЛИ -> Изпитен протокол
2. В дясната част на екрана се появява форма, която позволява да се въведат необходимите данни за създаването на изпитен протокол, (Фиг. 9).



Фиг. 9. Формуляр за разработка на изпитен протокол

ИС “Студент” автоматично генерира поредният номер на протокол и учебната година, в която за последно е бил записан студента. Стойностите на тези полета може да се коригират.

3. След натискане на бутона “Въведи”, под първоначално въведените данни се генерира протокол с всички студенти, отговарящи на зададените критерии.

4. Разпечатването на протокола се извършва с поставяне на отметката на “Печат” и натискане на бутона “Въведи”. На екрана ще се зареди протокола във вид, готов за разпечатване. За да се разпечата, се натиска бутона “Печат”.

Редактирането на готов изпитен протокол се извършва по следния начин:

1. От менюто се избира ПРОТОКОЛИ -> Изпитен протокол.

2. Попълват се полетата номер на протокол и учебна година. Натиска се бутона “Въведи” и ИС “Студент” ще изведе протокола, който е бил създаден с номера и учебната година, които са въведени.

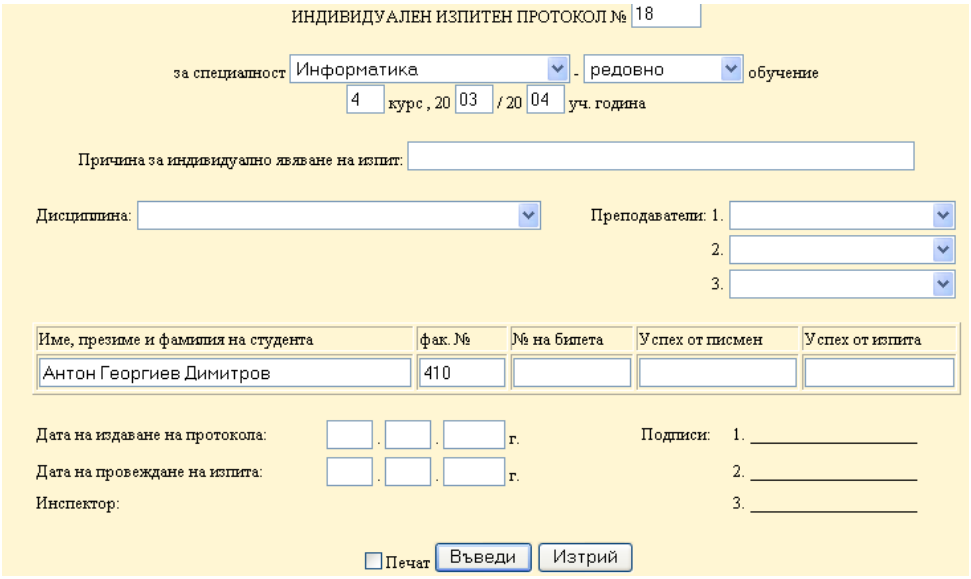
3. Разпечатването на протокола става както е описано в стъпки 4 и 5 от последователността за създаване на изпитен протокол.

РАЗРАБОТКА НА ИНДИВИДУАЛЕН ПРОТОКОЛ

За създаване на нов или редактиране на вече създаден индивидуален протокол, трябва да направите следното:

1. Избира се от менюто ПРОТОКОЛИ -> Индивидуален протокол и в дясната част на екрана се зареждат полета за въвеждане – „За нов протокол” и „За вече създаден протокол”.

Попълвайки полетата „За нов протокол” се визуализира формуляр за попълване на данни за конкретен протокол, (Фиг. 10).



Фиг. 10. Създаване на нов изпитен протокол

След попълване данните се натиска бутона “Въведи”. Ако е въведена некоректна информация, системата ще оцвети в червено полетата, които трябва да се коригират. Ако всички данни са коректно въведени – информацията ще се запише в БД.

Редактиране на вече създаден индивидуален протокол се извършва чрез попълване на полетата „За вече създаден протокол” и натискане на бутона “Въведи”. След като се актуализират данните се натиска бутона “Въведи”.

ГЛАВНА КНИГА

Оценките от изпитните и индивидуалните изпитни протоколи са част от главна книга. На тази основа са моделирани програмни части в ИС, които поддържат създаване на академическа справка, извлечение и уверение за съответния студент. Търсенето се извършва по ЕГН, фак. № и специалност. Предвидена е опция за печат на съответния формуляр.

ПРОТОКОЛ ЗА ДЪРЖАВЕН ИЗПИТ/ДИПЛОМНА РАБОТА

При полагане на държавен изпит или защита на дипломна работа се издава протокол. ИС “Студент” има възможност да генерира тези протоколи. За целта обаче трябва да се извършат няколко действия. За да се появят всички студенти, които ще полагат държавен изпит, трябва да се въведат данните им в съответните таблици в главната книга.

За да се яви студент на държавен изпит, той трябва да подаде молба. Щом той подаде молбата в съответния деканат, инспекторът по студентските въпроси въвежда данните в съответните таблици. За да се добави студент в протокол за защита на дипломна работа, задължително условие е да се попълни датата на предаване/депозиране на дипломната работа.

ЗАЩИТА НА ИНФОРМАЦИОННАТА СИСТЕМА „СТУДЕНТ”

ИС „Студент” е предназначена както за инспекторите по студентските въпроси, така и за преподавателите и студентите. При въвеждане на именника системата автоматично изпраща на електронната поща на студента потребителското му име и парола. Пароли и

потребителски имена често се забравят, това налага в ИС да има възможност за създаване, редактиране и изтриване на потребители. Това се извършва от менюто НАСТРОЙКИ -> Администриране. Там се създават потребителските акаунти за преподаватели, административни работници и студенти.

Моделирането на защитата на ИС „Студент" е извършено с използване на *p*-адичния псевдослучаен генератор [8], софтуерно изпълнен с клас `p_adic` [9].

ЗАКЛЮЧЕНИЕ

На основата на релационната база от данни „Студент" е разработена web базирана информационна система за нейното управление. Моделирани програмно са основните дейности извършвани от инспектора по студентските въпроси. Създаденото web приложение е лесно за работа и позволява непрекъсната актуализация.

ЛИТЕРАТУРА

1. **Ли, Дж., Б**. Уэр, *Использование Linux, Apache, MySQL и PHP Web-приложений*, ред. А. Н. Учниченко, Вильямс, 2004, 430 с.
2. **Маджаров**, И., *Интернет за персонални компютри. Web и базите данни*, 2001, 10 с.
3. ***Свитък шаблони от документи****, свързани с досието и академичната информация на студент в Шуменския университет „Епископ Константин Преславски"*, 2004.
4. **Стоянов**, Б., Д. Бояджиев, П. Милев, *Проектиране на релационна база от данни „Студент"*, Годишник на Шуменския университет „Епископ Константин Преславски", Факултет по математика и информатика, 2006, под печат.
5. **Стоянов**, Б., П. Милев, Р. Христова, *Модел на релационна база от данни „Недвижими имоти*", Годишник на Шуменския университет „Епископ Константин Преславски", Факултет по математика и информатика, Том XV С, 2002, с. 120-140.
6. ***MySQL Reference Manual***, `http://www.mysql.com`.
7. ***PHP Manual***, `http://www.php.net`.
8. **Stoyanov**, B., B. Bedzhev, *Algorithm for p-adic Combiner Generator Synthesis,* XXXIX International Scientific Conference on Information, Communication and Energy Systems and Technologies, ICEST 2004, 16-19 June, 2004, Bitola, Macedonia, pp. 345-348.
9. **Stoyanov**, B., B. Bedzhev, Zh. Zhekov, *Computation Model of p-adic Arithmetic,* XXXIX International Scientific Conference on Information, Communication and Energy Systems and Technologies, ICEST 2004, 16-19 June, 2004, Bitola, Macedonia, pp. 341-344.

ПРОЕКТИРАНЕ НА РЕЛАЦИОННА БАЗА ОТ ДАННИ „СТУДЕНТ”

Борислав П. Стоянов, Димо С. Бояджиев, Петър А. Милев

DESIGN OF RELATION DATABASE “STUDENT”

Borislav P. Stoyanov, Dimo S. Boyadzhiev, Petar A. Milev

Design of relation database “Student” is discussed in this paper. Main parts of design process are shown.

KEY WORDS: design of relation database, normalization process

УВОД

Проектирането на релационна база от данни е процес, който включва три основни фази: анализ на изискванията, моделиране на данните и нормализация [1], [3]. Спазвайки отделните фази е възможно получаването на устойчиви структури, поддържащи лесно модифициране на схемата си и съхраняваната в нея информация и лесно разработване на приложения за крайния потребител [8]. На тази основа, целта на настоящата статия е проектирането на релационна база от данни „Студент” за подпомагане дейността на инспектора по студентските въпроси към Шуменския университет „Епископ Константин Преславски”. За постигането на тази цел първо е направен анализ на дейността на инспектора по студентските въпроси и след това е моделирана и приведена в трета нормална форма релационна база от данни „Студент”. Извършена е софтуерна реализация на предложената структура. В заключение са направени изводи за получените резултати и са посочени насоките за по-нататъшната работа.

АНАЛИЗ НА ДЕЙНОСТТА НА ИНСПЕКТОРА ПО СТУДЕНТСКИТЕ ВЪПРОСИ

От запознаването с Правилници [4], [5], [6] и длъжностната характеристика на инспектора по студентските въпроси [2] на Шуменския университет „Епископ Константин Преславски” ставя ясно, че основният пакет документи поддържан за всеки студент и свързаната с него служебна документация [7] са на съхранение при инспектора по студентските въпроси и той отговаря за поддържането на данните. Въз основа на тази информация се правят справки, протоколи, заповеди, извлечения и др. Постоянната ръчна обработка на документите води до чести грешки, като неправилно попълнени име, ЕГН, адрес, семейно положение, учебен предмет и др. Възможно е при издаване на диплома за завършено висше образование тя да бъде с грешно име или с други съществени грешни данни, което да доведе до нейната невалидност. Подобен е случаят и с оценките от изпитите, които се нанасят от преподавателите в големите и неудобни главни книги. Контролът за недопускане на такива грешки е труден и продължителен и отнема ценно работно време.

Системата създава неудобства за инспекторите по студентските въпроси, които за всяка една справка разгръщат дебелите главни книги за да търсят в тях конкретни или група от студенти, да преписват или сравняват данните и чак след това да дописват и допечатват дадената заявка. Примери за това са издаването на „Академична справка” или „Статистическа справка за висшето образование”.

В първият случай, инспекторът трябва да намери всички учебни планове на студента, да ги въведе в компютъра и да разпечата бланка с всички дисциплини, хорариуми и

лекторите, които са заложени в плановете за обучение на студента. Следва преписване на оценките от главната книга върху бланката. Това отнема много време и нерви на служителя, а се налага да го прави за всички завършващи от съответния деканат, който той представлява.

Втората справка цели да покаже статистика за това, колко мъже и колко жени завършват, колко прекъсват, какъв процент са с успех отличен, какъв процент са социално слаби, процент семейни студенти и много други статистически данни, които за да се изчислят трябва да се извърши търсене във всички главни книги и всички студентски досиета. Това отнема ценно работно време на служителите. В следствие на това, инспекторите бързат и допускат грешки. Моделирането и нормализирането на база от данни свързана с този процес би подпомогнала неговата автоматизация, изключвайки в голяма степен човешката грешка и улеснявайки обслужването на студентите.

МОДЕЛИРАНЕ И НОРМАЛИЗИРАНЕ НА РЕЛАЦИОННА БАЗА ОТ ДАННИ „СТУДЕНТ”

След направеният анализ на студентската документация се предлагат следните 45 броя релационни таблици:

Потребители (#индекс, име, парола, e-mail, тип)

Академична справка (#номер, индекс, дата, забележка, декан, ректор)

Държавен изпит (#индекс, номер, дисциплина, форма, протокол, оценка, председател, номер на явяване, успех)

Държавни изпити (#код, дисциплина)

Протокол за дипломна дисциплина (#номер, катедра, специалност, форма, магистратура, степен, учебна година, индекс, тема, успех, дата, комисия)

Дипломна работа (#индекс, номер, дата на вземане, тема, ръководител, дата на предаване, оценка, завеждащ катедра, номер на протокол, дата на протокол, председател)

Дисциплини (#код, дисциплина)

Държавен протокол (#номер, дисциплина, специалност, форма, степен, учебна година, индекс, явяване, успех, дата, комисия)

Индивидуален план (#индекс, одобрил, отхвърлил, учебна година, име, специалност, форма, курс, фак. номер, степен, дисциплина, семестри, лекции, упражнения, лектор, професионална квалификация)

Индивидуален протокол (#номер, индекс, специалност, форма, курс, учебна година, причина, дисциплина, лектор, номер билет, вид изпит, дата издаване, дата на изпита)

Изходни данни (#индекс, серия на дипломата, номер на дипломната, регистрационен номер, дата на регистриране, научна степен, специалност, специализация, професионална квалификация, дата на получаване на дипломата, дата - архив)

Катедри (#код, катедра)

Деца (#ЕГН, име)

Имена (#ЕГН, първо име, презиме, фамилия)

Снимки (#ЕГН, адрес)

Лична карта (#ЕГН, номер, дата на издаване, място на издаване)

Настоящ адрес (#ЕГН, място, пощенски код, община, област, улица, улица номер, квартал, блок, вход, етаж, апартамент, телефон)

Професионална квалификация (#код, квалификация)

Изпитен протокол (#номер на протокол, специалност, форма, индекс, курс, учебна година, дисциплина, код на преподавател, код на преподавател 2, писмен- редовна, изпит- редовна, писмен- поправка, изпит- поправка, дата- редовна, дата- поправка, писмен- ликвидация, изпит- ликвидация, дата- ликвидация)

Постоянен адрес (#ЕГН, пощенски код, община, област, улица, номер улица, квартал, блок, вход, етаж, апартамент, телефон)

Общи данни (#ЕГН, соц. произход, семейно положение, военна служба, гражданство)

Раждане (#ЕГН, град, община, област, страна)

Родители (#ЕГН, номер, вид на роднината, име, жизнен статус)

Адрес на родители (#ЕГН, номер, град, община, област, улица, номер улица, квартал, блок, вход, етаж, апартамент)

Семейство(#ЕГН, вид роднина, име)

Средно образование (#ЕГН, серия, регистрационен номер, дата, среден успех, училище, град, тип на образованието, име на училището)

Трудов стаж (#ЕГН, къде, кога)

Трудов стаж сега (#ЕГН, къде, кога)

Предишно висше образование (#ЕГН, име, вид, град, специалност)

Магистърски програми (#код, име)

Научна степен (#код, научна степен)

Учебен план (#код, дата на приемане, специалност, вид обучение, срок, научна степен, дата на създаване)

Преподаватели (#код, преподавател)

Оценки (#индекс, курс, учебна година, дисциплина, номер на явяване, номер на протокол, дата на явяване, код на преподавател, тип)

Практики и стажове (#индекс, номер, тема, град, от дата, до дата, брой седмици, оценка)

Предишен ВУЗ (#индекс, дата на записване, специалност, форма, заповед, заповед номер, бал, научна степен)

Специалности (#код, специалност)

Статус на студента (#код, статус)

Студентско положение (#индекс, курс, учебна година, дата зимен, среден успех от зимен, заверка зимен, дата летен, среден успех летен, заверка летен, прекъснал/ презаписал, среден годишен успех)

Тип дисциплини (#код, тип)

Учебен план (#код, професионална квалификация, тип дисциплина, дисциплина, лекции(общо), упражнения (общо), лекции, семинарни, лабораторни, оценка, семестър, курс, код на преподавателят)

Учебни планове (#код, тип изпит, забележка, декан, ректор)

Уверение (#номер, индекс, дата на издаване, полугодие, месец на завършване, година на завършване, свободен текст, "да послужи пред")

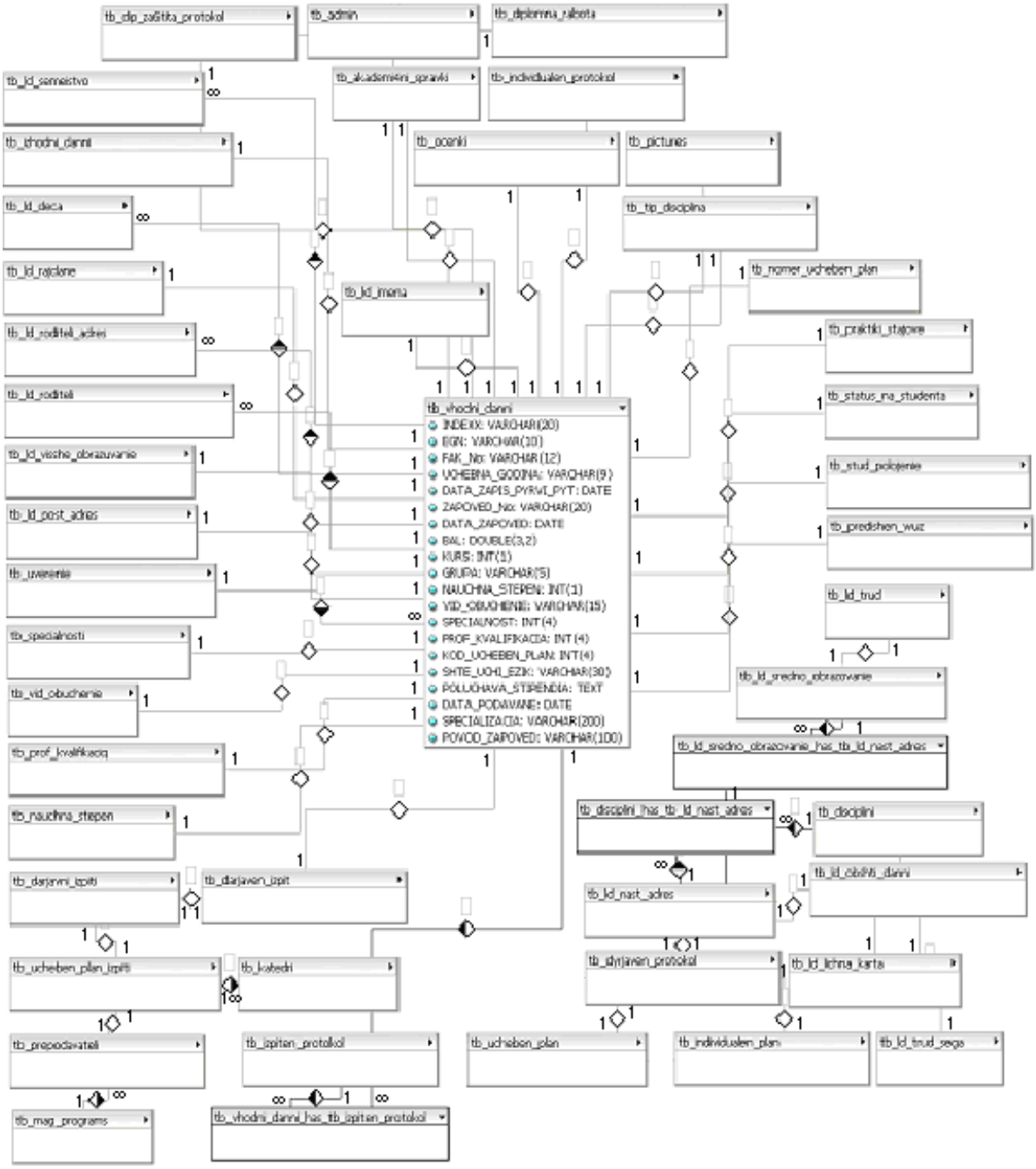
Входни данни (#индекс, ЕГН, Фак. номер, учебна година, дата на която е записан за пръв път, заповед номер, дата на заповедта, бал, курс, група, научна степен, вид обучение, специалност)

Вид обучение (#код, вид)

Информацията, която може да се запазва в тях напълно моделира информационните единици, свързани с работата на инспектора по студентските въпроси. Възможна е пълната автоматизация на работния процес.

Релационната схема „Студент” е нормализирана до 3 нормална форма. По този начин са избегнати загубата и повторението на информация, аномалиите при актуализация на данните, отстранени са нежелателните функционални зависимости между атрибутите на релационните схеми. Нормализирането е направено по време на процеса на моделиране.

Моделирането е извършено визуално с помощта на системата за проектиране на бази от данни DBDesigner 4 [9]. Нормализираната схема е показана на Фиг. 1.



Фиг. 1. Модел на релационна база от данни „Студент” в 3 нормална форма

Проектираният визуален модел на релационна база от данни „Студент” е експортиран в SQL-скриптов файл. Този файл е изпълнен в MySQL командния интерпретатор [10]. В

резултат в MySQL сървъра за бази от данни е генерирана релационната база от данни „Студент” съдържаща всички предложени таблици.

ЗАКЛЮЧЕНИЕ

С цел проектиране на релационна база от данни „Студент” в настоящата статия е анализирана дейността на инспектора по студентските въпроси към Шуменския университет „Епископ Константин Преславски”. В съответствие с основните информационни единици, които участват в процеса на обслужване на студентите е моделирана текстово и графично релационна база от данни „Студент”, приведена в 3 нормална форма.

ВИЖДАНИЯ ЗА НАСОКИТЕ НА ПО-НАТАТЪШНАТА РАБОТА

Насоките на по-нататъшната работа може да се обобщят в следните направления: продължаване работа по увеличаване пълнотата на релационната база от данни „Студент”и създаване на софтуерно приложение за работа с нея.

ЛИТЕРАТУРА

1. **Азълов**, П., *Бази от данни, релационен и обектен подход*, Техника, София, 1991, 208 с.1
2. *Длъжностна характеристика на инспектор по студентските въпроси в Шуменския университет „Епископ Константин Преславски”*, 2004.7
3. **Милев**, П., Р. Христова, В. Дянкова, *Бази от данни и приложения*, Университетско издателство, 2002, 361 с.2
4. ***Правилник*** *за приемане на студенти в Шуменския университет „Епископ Константин Преславски”*, Университетско издателство, 2004.4
5. ***Правилник*** *за структурата и организацията на учебния процес в Шуменския университет „Епископ Константин Преславски”*, Университетско издателство, 2004.5
6. ***Правилник*** *за устройството и дейността на Шуменския университет „Епископ Константин Преславски”*, Университетско издателство, Шумен, 2004, 39 с.6
7. ***Свитък шаблони*** *от документи, свързани с досието и академичната информация на студент в Шуменския университет „Епископ Константин Преславски”*, 2004.8
8. **Стоянов**, Б., П. Милев, Р. Христова, *Модел на релационна база от данни „Недвижими имоти*”, Годишник на Шуменския университет „Епископ Константин Преславски”, Факултет по математика и информатика, Том XV С, 2002, с. 120-140.3
9. ***DBDesigner 4***, *Online Dokumentation*, version 1.0.42, `http://www.fabforce.net`.9
10. ***MySQL Reference Manual***, `http://www.mysql.com`.10

ОСНОВНИ КРИПТОГРАФСКИ СВОЙСТВА НА ПСЕВДОСЛУЧАЙНИТЕ ГЕНЕРАТОРИ НА РЕДИЦИ

БОРИСЛАВ П. СТОЯНОВ

BASE CRYPTOGRAPHIC PROPERTIES OF THE PSEUDORANDOM GENERATORS OF SEQUENCES

BORISLAV P. STOYANOV

This article presents four basic cryptographic properties of the pseudorandom generators of sequences – period, linear complexity, speed and statistical properties.

KEY WORDS: cryptography, pseudorandom generators

ВЪВЕДЕНИЕ В ПРОБЛЕМА

Псевдослучайните генератори (ПСГ) са основните изграждащи елементи в голяма част от разпространените в момента поточни и блокови шифри. За да могат да бъдат използвани като такива трябва да бъдат изследвани определени техни характеристики [1], [2], [4], [5], [7], [10], [12], [13], [14], [17], [21], [22], [25], [27], [32], [34], [35], [36], [37], [39], [41] - период и линейна сложност на псевдослучайната им редица, скорост на генериране на изходната гама и да бъдат извършени статистически изследвания. Целта на тази статия е да се анализира криптографската важност на тези характеристики и да се приведат данни за популярни в момента псевдослучайни генератори.

ОСНОВНИ СВОЙСТВА НА ПСЕВДОСЛУЧАЙНИТЕ ГЕНЕРАТОРИ НА РЕДИЦИ

В този параграф се обсъждат периода, линейната сложност, скоростта и статистическите изследвания – основните свойства характеризиращи псевдослучайните генератори.

ПЕРИОД И ЛИНЕЙНА СЛОЖНОСТ

Много важна характеристика на псевдослучайната редица е дължината на периода. Това е така, защото най-често генераторът на ключов поток е с краен брой вътрешни състояния. Практически, дължина на псевдослучайна редица над 2^{128} бита е напълно достатъчна за да осигури избягване едни и същи нейни части да се използват два пъти по време на шифрирането.

Колкото е по-дълъг периодът толкова е по-голяма линейната сложност на псевдослучайната редица. Необходимостта от висока линейна сложност е доказана в една от значимите публикации свързани с анализа на поточните шифри дължаща се на Massey [30]. В нея е описан алгоритъмът на Berlekamp-Massey, установяващ линейната сложност на редица. Той е криптографски важен, тъй като линейната рекурентност, удовлетворена чрез редица с линейна сложност λ, може ефективно да се изчисли въз основа на 2λ последователни бита от редицата. Висока линейна сложност означава, че голяма част от редицата трябва да се наблюдава и че се изисква много дълъг линеен преместващ регистър с обратна връзка за да се преповтори тази редица. Проверката за достатъчност на линейната

сложност се извършва обикновено с помощта на статистически тест, например с пакет NIST test suite [38], [40] (*National Institute of Standards and Technologies*, Национален институт по стандарти и технологии).

СКОРОСТ

Една от определящите характеристики за избор на един пред друг криптографски алгоритъм е скоростта на генерация на изходната гама. Измерванията се извършват в MByte/Sec. Общоприето е мнението, че скорост между 1 и 10 MByte/Sec е напълно достатъчна за една криптосистема [18]. В Табл. 1 са показани изходните скорости на следните популярни в момента псевдослучайни генератори: линеен, сравним по модул ПСГ (*linear congruential, LC*), кубичен, сравним по модул ПСГ (*cubic congruential, CC*), Blum-Blum-Shub ПСГ (*BBS*) и 3DES (*Data Encryption Standart*) ПСГ [32], [39] [44]:

Табл. 1. Скорост на генериране на изходната гама на LC, BBS, CC и 3DES ПСГ-и

Генератор	**Скорост(MByte/Sec)**
LC	0.031
BBS	0.070
CC	1.004
3DES	9.848

Компютърната конфигурация, на която са направени измерванията на LC и CC генератори (AMD Athlon™ XP 2200+ 1.81 GHz, 256 Mbytes RAM под Windows XP SP 2) е подобна по производителност на компютърната конфигурация използвана за измерване на останалите генератори [44]. Всички C++ модули са оптимизирани за скорост, имат включени преизчислими таблици и 386 ASM модули за събиране и изваждане с плаваща запетая. Забелязва се неравномерност в получените параметри поради различната вътрешна структура на псевдослучайните генератори и разликите в софтуерното им изпълнение.

СТАТИСТИЧЕСКИ ИЗСЛЕДВАНИЯ

Ясно е, че периодът и линейната сложност , за които е препоръчително да бъде изследвана всяка псевдослучайна редица, сами по себе си не са достатъчни да обхванат всички характеристики на случайно изглеждащата редица за да се докаже криптографската й сигурност. Затова е необходимо да се приложи широк набор от различни статистически тестове, за да се оцени до колко вярно е твърдението, че е била генерирана перфектна случайна редица [16].

Съществуват няколко общопризнати източници на статистически тестове, които предлагат изследване на псевдослучайни битови редици и емпирично доказателство в подкрепа на тяхната случайност. Класически списък от 10 статистически теста е даден от Knuth [6]. Тестовият пакет DIEHARD, разработен от Marsaglia [29], съдържа набор от 15 теста. Този комплект включва тестове, които са по-строги в изискванията си към входните редици, отколкото тези на Knuth, в смисъл че много популярни генератори пропадат при тестване с него. Стандартът FIPS 140-1 [19] и поправката му FIPS 140-2 [20] на САЩ (*Federal Information Processing Standard*, Федерален стандарт за информационно развитие) предлага описание на 4 статистически теста, а софтуерната библиотека Crypt-XS [23], разработена от Технологичния университет в Куийнсленд, Австралия, предлага 6 такива

теста. Друг голям тестов пакет е т.н. TestU01 [28], разработка на L'Ecuyer и Simard от Universite de Montreal.

В резултат на съвместната работа между отдел „Компютърна сигурност" и „Статистическо инженерство" на NIST в САЩ по-известните и криптографски важни статистически тестове от изброените по-горе пакети са обединени и доразвити в NIST test suite. Това са тестовете - честотен тест, блоков честотен тест, тест за редици, тест за най-дългата редица от единици в блок, двоичен тест матричен ранк, спектрален тест на основата на дискретното преобразуване на Фурие, тест за незастъпващи се непериодични шаблони, тест за застъпващи се непериодични шаблони, универсален статистически тест на Maurer [31], тест за Lempel-Ziv компресия [45], [46], тест за линейна сложност, сериен тест [33], приблизителен ентропиен тест, тест за натрупващи се (кумулативни) суми, тест „случайна разходка" и тест „случайна разходка"-вариант. След публикация на Kim, Umeno и Hasegava [26], поради открити грешки, спектралният тест е коригиран, а Lempel-Ziv тестът е изключен от NIST test suite.

Посредством множеството от 15 различни функционални статистически теста на NIST test suite, се определя дали свойствата на псевдослучайната битова редица са близки до случайните. Той е полезен в следните насоки:

- установява, кои ПСГ-и произвеждат криптографски слаби двоични редици;
- спомага за по-добрия дизайн на нови ПСГ-и;
- потвърждава коректното софтуерно или хардуерно изпълнението на ПСГ;
- изучава съществуващи вече ПСГ-и;
- установява степента на случайност на текущо използвани ПСГ-и.

Този тестов пакет може определи следните свойства, които притежават редиците, генерирани от ПСГ:

1. *Равномерна разпределеност*; т.е. дали вероятността от появяване на единица или нула е еднаква. Нулите и единиците в подблоковете на редицата трябва да са приблизително едни и същи както в една истинска случайна редица (честотен и блоков честотен тест). Броят от редици от нули и единици с различни дължини трябва да е такъв, какъвто се очаква в една случайна редица (тест за редици). Не трябва да съществуват линейни зависимости между подредици от тестваната редица (тест за най-дългата редица от единици в блок). Освен това всеки $m-$битов шаблон трябва да има същата вероятност да се появи, както всеки друг $m-$битов шаблон (сериен тест).

2. *Мащабируемост*, т.е. всеки тест приложен към дадена редица може да бъде приложен и към произволно взети от нея подредици. Броят на срещанията на незастъпващи се и застъпващи се непериодични шаблони в редицата следва да е в определени граници (тестове за незастъпващи се и застъпващи се непериодични шаблони). Честотата на поява на припокриващи се блокове с последователни дължини трябва да е същата както при случайна редица (приблизителен ентропиен тест). Броят на посещенията до отделни битове в тестваната редица следва да е под определена граница (тестове случайна разходка и случайна разходка-вариант).

3. *Некомпресируемост*, т.е. не трябва да е възможно редицата да бъде значително компресирана без загуба на информация (универсален тест на Maurer). Натрупващата се сума от срещащите се подредици трябва да е както тази в истинска случайна редица (тест натрупващи се суми).

4. *Неповторяемост*, т.е. редицата не трябва да съдържа повтарящи се, близки един до друг шаблони (спектрален тест). Линейната сложност следва да е достатъчна за да неможе редицата да се преповтори с прекалено къс линеен преместващ регистър с обратна връзка (тест линейна сложност).

Изследването на псевдослучайни генератори би било непълно без съпоставянето им със съществуващите и вече утвърдени псевдослучайни генератори. Резултатите от тестване с NIST test suite на LC, CC, BBS и 3DES са показани в Табл. 2.

Табл. 2. Статистическо изследване с NIST test suite на LC, BBS, CC и 3DES ПСГ-и

	NIST test suite	
	Пропорция на преминалите теста редици в %	***Разпределение на всички P-value в %***
LC	100	93
CC	73	67
BBS	100	100
3DES	100	100

В два от псевдослучайните генератори се наблюдават слабости в изходните редици. Това са линейният, сравним по модул ПСГ, който показва лошо разпределение на стойностите в спектралния тест, а кубичният, сравним по модул ПСГ се проваля в голяма част от проведените тестове.

Добрите резултати, показани при статистическо тестване с NIST test suite, са достатъчни ПСГ да се приеме за генератор произвеждащ псевдослучайни редици с качества, близки до тези на истинска случайна редица. Това е така поради факта, че за голяма част от ПСГ-и, които се използват не са известни други методи за оценка на случайността на ключовата им гама [15], [24], [36], [39].

ЗАКЛЮЧЕНИЕ

Анализирани се основните свойства – период, линейна сложност, скорост и статистически данни, които се изследват при всеки псевдослучаен генератор използван за защита на информация. Приведени са резултати за скоростта и статистическите свойства, изследвани с NIST test suite, на LC, BBS, CC и 3DES ПСГ-и.

Дискутираните криптографски свойства са използвани за успешен дизайн и тестване на *p*-адични комбиниращи псевдослучайни генератори на редици [3], [8], [9], [11], [42], [43].

НАСОКИ НА ПО-НАТАТЪШНАТА РАБОТА

Насоките на по-нататъшната работа включват разширяване броя на изследваните криптографските свойства чрез изследване важността на прилаганите криптографски атаки срещу различни псевдослучайни генератори и анализиране на съществуващи генератори с разглежданите свойства.

ЛИТЕРАТУРА

1. **Антонов**, П., С. Малчев, *Криптография в компютърните комуникации*, Варна, 2000, 315 с.
2. **Баричев**, С., Р. Серов, *Основы современной криптографии*, version 1.2, Горячая линия – Телеком, Москва, 2001, 122 с.
3. **Беджев**, Б., Ж. Ташева, Б. Стоянов, В. Матеев, *Автокорелационни свойства на сигнали, формирани от свиващ-мултиплексиращ генератор*, Трудове на научната сесия на РУ „Ангел Кънчев”, 2004, 121-126 с.

4. **Берлекэмп**, Э., *Алгебраическая теория кодирования*, Перевод с английского И. Грушко, Москва, Мир, 1971, 477 с.
5. **Додунеков**, Ст., Й. Денев, *Кодиране на информацията*, Народна просвета, София, 1985, 138 с.
6. **Кнут**, Д., *Искусство программирования для ЭВМ, т. 2, Получисленные алгоритмы*, Перевод с английского, Под редакцией Бабенко, К. И, Москва, Мир, 1977, 724 с., http://www-cs-faculty.stanford.edu/~knuth/taocp.html.
7. **Петров**, Р., *Защита на информацията в компютрите и мрежите*, София, Корени, 2002, 320 с.
8. **Ташева**, Ж., *Изследване на свойствата на свиващ-мултиплексиращ генератор чрез приблизителен ентропиен тест*, Трудове на научната сесия на РУ „Ангел Кънчев”, 2004, 115-120 с.
9. **Ташева**, Ж., Б. Стоянов, *Спектрален анализ на свиващ-мултиплексиращ генератор*, Научна сесия на НВУ „В. Левски”, Част II, 2004, 105-112 с.
10. **Цирлер**, Н. *Линейные возвратные последовательности*, Перевод с английского, Кибернетический сборник, No. 6, ИЛ, Москва, 1963, 55-79 с.
11. **Bedzhev**, B., Zh. Tasheva, B. Stoyanov, *Summation-Shrinking Generator*, International Conference “Information Technologies and Security”, ITS 2004, 22-26 June, 2004, Partenit, Crimea, Ukraine, pp. 119-127.
12. **Beker**, H., F. Piper, *Cipher Systems, The Protection of Communications*, New York: van Nostrand Reinhold, 1982, p. 427.
13. **Bellare**, M., P. Rogaway, *Introduction to Modern Cryptography*, November 24, 2001, p. 203, http://www-cse.ucsd.edu/users/mihir.
14. **Blum**, L., M. Blum, M. Shub, *A Simple Unpredictable Pseudo-Random Number Generator*, SIAM Journal of Computing, Vol. 15, No. 2, 1986, pp. 364–383.
15. **Burr**, W., *Selecting the Advanced Encryption Standard*, IEEE Security & Privacy, March/April 2003, pp. 43-52, http://computer.org/security/.
16. **Chess**, B., M. Gary, *Static Analysis for Security*, IEEE Security & Privacy, November/December 2004, pp. 76-79, http://www.computer.org/security/.
17. **Cusick** T., C. Ding, A. Renvall, *Stream Ciphers and Number Theory*, Revised Version, North-Holland Mathematical Library, Vol. 55, August 2003, p. 474.
18. **Damgard**, I., *A Quick Introduction to some Crypto Concepts*, PDS 2002, February 18, 2002, p. 18.
19. **FIPS 140-1**, *Security Requirements for Cryptographic Modules,* Federal Information Processing Standards Publication 140-1, U.S. Department of Commerce / NIST, National Institute of Standards and Technology, Springfield, Virginia, January 11, 1994, p. 53.
20. **FIPS 140-2**, *Security Requirements for Cryptographic Modules,* Federal Information Processing Standards Publication 140-2, U.S. Department of Commerce / NIST, National Institute of Standards and Technology, May 25, 2001, p. 64.
21. **Goldwasser**, S., M. Bellare, *Lecture Notes on Cryptography*, August 2001, p. 283, http://theory.lcs.mit.edu/~shafi.
22. **Golomb**, S., *Shift Register Sequences*, Holden-Day, San Francisco, 1967, Reprinted by Aegean Park Press, Laguna Hills, CA, USA, 1981, p. 257.
23. **Gustafson**, H., et. al., *A computer package for measuring strength of encryption algorithms*, Journal of Computers and Security, Vol. 13, No. 8, 1994, Crypt-XS Suite of Statistical Tests, pp. 687-697, http://www.isrc.qut.edu.au/cryptx/index.html.
24. **Hellekalek**, P., S. Wegenkittl, *Empirical Evidence Concerning AES*, Preprint accepted by the ACM Transactions on Modeling and Computer Simulation, p. 12, 2003.

25. **Helleseth**, T., P. Kumar, *Mobile Communications Handbook. Chapter 8. Pseudonoise Sequences*, Ed. Suthan S. Suthersan, Boca Raton, CRC Press LLC, 1999, p. 14.
26. **Kim**, S., K. Umeno, A. Hasegawa, *Corrections of the NIST Statistical Test Suite for Randomness*, Cryptology ePrint Archive, Report 2004/018, 2004, p. 14, http://eprint.iacr.org/2004/018.
27. **Klapper**, A., M. Goresky, *Large Period Nearly de Bruijn FCSR Sequences, Advances in Cryptology*, EUROCRYPT '95, LNCS Vol. 921, Sringer-Verlag, N. Y., 1995, pp. 263-273.
28. **L'Ecuyer**, P., R. Simard, *TestU01: A Software Library in ANSI C for Empirical Testing of Random Number Generators. Software User's Guide*, Department d'Informatique et de Recherche Operationelle, Universite de Montreal, 2002.
29. **Marsaglia**, G., *DIEHARD: a battery of tests of randomness*, 1996, http://stat.fsu.edu/~geo/diehard.html.
30. **Massey**, J., *Shift Register Synthesis and BCH Decoding*, IEEE Transactions on Information Theory, Vol. 15, 1969, pp. 122-127.
31. **Maurer**, U., A *Universal Statistical Test for Random Bit Generators*, Journal of Cryptology, Vol. 5, 1992, pp. 89–105.
32. **Menezes**, A., P. van Oorshot, S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1997, p. 780, http://www.cacr.math.uwaterloo.ca/hac.
33. **Niederreiter**, H., *The Serial Test for Pseudorandom Numbers Generated by the Linear Congruential Method*, Numer. Math. 46 (1985), pp. 51–68.
34. **Paar**, C., *Applied Cryptography and Data Security. Lecture Notes (version 2.5 – January 2005)*, Department of Electrical Engineering and Information Sciences, Ruhr-Universitat Bochum, Germany, 2005, p. 198, http://www.crypto.rub.de.
35. **Reeds**, J., N. Sloane, *Shift Register Synthesis (modulo m)*, SIAM Journal of Computing, August 1985, pp. 505-513.
36. **Robshaw**, M., *Stream Ciphers*, RSA Laboratories Technical Report TR-701, Version 2.0, July 25, 1995, p. 42.
37. **Rueppel**, R., *Analysis and Design of Stream Ciphers*, Springer Verlag, Berlin, 1986, p. 244.
38. **Rukhin**, A., J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray, S. Vo. *A Statistical Test Suite for Random and Pseudo-Random Number Generators for Cryptographic Application*, NIST Special Publication 800-22 (with revision May 15, 2001), p. 162.
39. **Schneier**, B., *Applied Cryptography*, 2nd Edition, John Wiley and Sons, New York, 1996, p. 758.
40. **Soto**, J., *Statistical Testing of Random Number Generators*, NIST Special Publication, p. 12, http://csrc.nist.gov/rng/.
41. **Stinson**, D., *Cryptography: Theory and Practice*, 2nd Edition, CRC Press, 2002, p. 434.
42. **Stoyanov**, B., *2-adic Summation-Shrinking Generator*, Western European Workshop on Research in Cryptology, WEWORC 2005, Leuven, Belgium, 5-7 July, 2005, pp. 103-104.
43. **Tasheva**, Zh., B. Bedzhev, B. Stoyanov, *P-adic Shrinking-Multiplexing Generator*, IEEE Third International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2005, Sofia, Bulgaria, 5-7 September, 2005, pp. 443-448.
44. **Wei**, D., *CRYPTO++ 5.2.1 Benchmarks*, 23 July, 2004, p. 6., http://www.eskimo.com/~weidai/benchmarks.html.
45. **Ziv**, J., A. Lempel, *A Universal Algorithm for Sequential Data Compression*, IEEE Transactions on Information Theory, Vol. 23, No. 3, May 1977, pp. 337-343.
46. **Ziv**, J., *Compression, Tests for Randomness and Estimating the Statistical Model of An Individual Sequence*, Sequences (ed. R. Capocelli), Berlin, Springer-Verlag, 1990, pp. 530-536.

ЧИСЛОВИ ФОРМАТИ ОТ ТИП ДАТА И ВРЕМЕ В ЕЛЕКТРОННИ ТАБЛИЦИ

ТЕОДОСИ К. ТЕОДОСИЕВ, ЕЛЕНА Т. ДИМИТРОВА

NUMBER FORMATS OF TYPE DATE&TIME IN SPREADSHEET

THEODOSI K. THEODOSIEV, ELENA T. DIMITROVA

This paper describes an interest part of the module "Spreadsheet" in the program of "Information technologies". Processing with the different type of data in particular with Date&Time is hard to put it for students but on other side this problem isn't look in detail in the books. There we share some ideas and tasks for surmount difficulties in study of specified number type.

KEY WORDS: Teaching, Spreadsheet, Number type, Date&Time

1. Въведение. Предлаганата работа разглежда интересна част от модула **„**Електронни таблици" в програмата по дисциплината „Информационни технологии", изучавана в училищата. Обработването на различни по тип данни и в частност с типа ДАТА/ВРЕМЕ затруднява обучаемите, а от друга страна не е подробно разгледан в учебниците. Тук споделяме някои идеи и задачи за преодоляване на трудностите при изучаването на посочения числов тип. Във **глава 2** се коментират целите на обучението в модула „Електронни таблици" и основните понятия и дейности свързани с табличното представяне на информацията. В **глава 3** са разгледани особеностите на числовите формати от тип Date и Time и използването им в таблиците. Интересни примери за използване на вградените функции за този тип данни са разгледани в **глава 4** и са коментирани някои методични бележки. В **глава 5** са дадени още няколко практически задачи, чрез които се дава пример за моделиране чрез табличен процесор. Обърнато е внимание на допълнителните полета и колони и използването на абсолютен адрес във формулите.

2. Цели на обучението по Електронни таблици. Овладяването на програми за създаване и обработка на електронни таблици се превърна в задължителна част от компютърната грамотност на съвременния образован човек. Именно затова те са включени като отделен модул в програмата по дисциплината „Информационни технологии", изучавана в училищата.

Усвояването на знания за табличния процесор и умения за създаване на електронна таблица е невъзможно без познаването на начините за изчисления, посредством аритметични операции, формули и използването на вградените в продукта функции. Целите на обучението са:

- Въвеждане и усвояване на аритметичните операции;
- Линейната форма на записване на математическите формули;
- Усвояване на приоритета при изпълнението на отделните операции;
- Възможностите за визуализация, редактиране и преизчисляване на формулите в електронната таблица.

Усвояването на учебното съдържание гарантира решаването на стандартни таблични задачи самостоятелно от учениците. В процеса на разрешаване на конкретни ситуации, последователно във времето, систематизирано се въвеждат и усвояват основните понятия и дейности, свързани с табличното представяне, въвеждане, редактиране, оформяне, съхраняване, обработка и извеждане на информацията. Потребността от автоматизирано извършване на разгледаните дейности се използва за мотивирано представяне и усвояване на

съответните инструментални средства на използвания табличен процесор. Едно от големите предимства на тези програми е възможността им да обработват различни по тип данни, което от друга страна поражда проблеми при изучаването и осмислянето им от учениците. От преподавателския си опит и от многобройните дискусии, които сме водили с колеги по проблемите на обучението по Електронни таблици, се налага извода, че обучаемите срещат трудности при осмислянето и обработването на различни по тип данни и в частност с типа ДАТА/ВРЕМЕ, който е предмет на настоящата публикация. Нейната цел е скромен опит да се споделят някои идеи и задачи за преодоляване на трудностите при изучаването на посочения числов тип.

3. Числови формати от тип Date и Time. Програмният продукт **Excel** работи с два основни типа данни- числови и символни. Това могат да бъдат числа, дати, часови стойности, текст, логически стойности, формули или стойности за грешки. Тяхното изобразяване в електронната таблица съответства на стандартни или избрани от потребителя формати, които се поддържат от програмата.

Работата със стойности от тип **Date&Time** (дати и часове) по своята същност е работа с числа. Интерес представляват задачи, в които се правят изчисления с данни от този тип. За съжаление при обучението в модула "Електронни таблици", на тези типове не се обръща нужното внимание. Те почти не се споменават в учебниците по „Информационни технологии" за 9 и 10 клас, а възможността за извършване на изчисления с различни по тип числови стойности е едно от най-големите предимства на табличния процесор. Наблюденията показват, че затруднения се срещат при конструиране на формули, в които участват клетки, съдържащи данни от тип **Date&Time**. За успешно справяне с проблема е необходимо усвояването, разширяването и задълбочаването на знанията, свързани с представянето на числовите стойности в електронни таблици. Целта е, да се формират умения и компетенции за използване на числови стойности от тип "дата" и "време" в таблични задачи. Към това спадат:

- Формати за представяне на дати и часови стойности;
- Въвеждане на серия от дати, дни на седмицата, месеци от годината или години в област от клетки от електронната таблица (фиг.1.1);
- Изчисления със стойности от тип дата и време;
- Вградени в табличния процесор функции за работа с посочените типове;
- Комбиниране на горните умения;

	A
1	14.12.2005
2	16.12.2005
3	18.12.2005
4	20.12.2005
5	22.12.2005
6	24.12.2005
7	26.12.2005

C
12:30
12:45
13:00
13:15
13:30
13:45
14:00

Фиг.1.1

Учениците трябва да усвоят форматите за представяне на дати и часови стойности, да осъзнаят същността и приложението на изчисленията с тях. Нужно е да се акцентира и върху това как се съхраняват датите и часовете от самата програма при въвеждането им от потребителя – като числа:

- Дати са последователни числа (серийни номера), които показват колко дни са изминали от някаква начална дата;
- Часът се съхранява като десетична дроб, която показва изминалата част от 24-часов период;

Според формата, по който се въвеждат, програмата ги разпознава и определя като числови стойности или текст (табл.1)

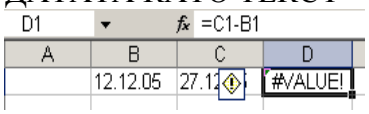
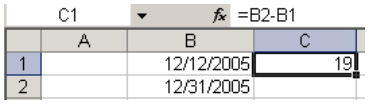
ПРИМЕР	ДЕЙСТВИЕ
ДАТАТА КАТО ТЕКСТ 	Така въведените данни са от текстов тип (вижда се и от подравняването) и резултата от формулата е съобщение за грешка.
ДАТАТА КАТО ТИП ДАТА/ВРЕМЕ 	В този случай е възможно да се изчисли интервала от време между двете дати, защото те са въведени във формат, позволяващ разглеждането им като числови стойности.

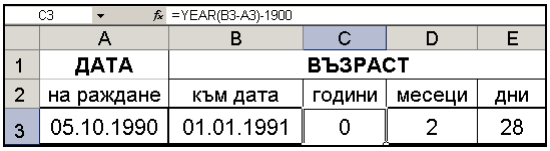
табл. 1

Следователно датите и часовете се въвеждат като стойности в клетките, за да се използват във формули при изчисления.

4. Вградени функции за данни от тип Date&Time. Необходимо и познаване на вградените функции за извличане на номера на деня, месеца, годината, на часа, минутите и секундите от времева числова стойност. Всичко това е възможно чрез подбор на подходящи задачи, които да се представят за решаване пред обучаемите.

Използването на данни от тип **Date** при работа с електронни таблици е целта на следните задачи:

1 зад. По зададена дата на раждане и текуща дата да се изчисли възрастта на човек с точност до 1 ден, чрез използване на таблица.



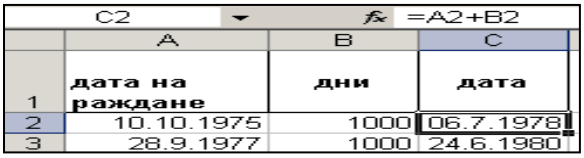
фиг. 1.2

На фигурата се вижда функцията за изчисляване на годините. За определяне на броя на месеците и дните се прилагат следните формули съответно в клетките D3 и E3:

За месеци - = MONTH(B3-A3)-1, а за дните =DAY(B3-A3).

2 зад. Да се състави таблица, чрез която да се изчислява по дадена дата на раждане на коя дата човек ще бъде на 1000 дни.

Тук учениците трябва да обработват датите, като имат предвид представянето им като серийни номера. За колоната с дати трябва да се посочи формат **Date**, а за броя дни **NUMBER** без цифри след десетичната точка. За новата дата отново се задава тип **Date** и се въвежда формулата: **= A2+B2**, ако съответните колони са A и B(фиг. 1.3).



Фиг. 1.3

Изграждането у учениците на умения за изчисляване на времеви интервали е цел на следващите задачи.

3 зад. Да се състави таблица за изчисляване времето и средната скорост на всеки участник в пробег от 25 км. на колоездачно състезание.

4 зад. Да се състави таблица за изчисляване на времето и цената, която трябва да заплати клиент на компютърна зала, ползвал услугите й.

Важно тук е изчисления интервал от време да се превърне в часове. Двете задачи илюстрират операции с данни от тип "време" и изискват прилагане на съответните за типа вградени функции.

F3 = (HOUR(D3)+MINUTE(D3)/60+SECOND(D3)/3600)

	A	B	C	D	E	F
1			пробег	/в км./	25	
2	**участник**	**старт**	**финал**	**време**	**ср.скорост**	**време в ч.**
3	1	15:00:00	17:21:09	02:21:09	10,63	2,3525
4	2	15:00:00	19:35:09	04:35:09	5,45	4,58583333
5	3	15:00:00	17:24:43	02:24:43	10,37	2,41194444
6	4	15:00:00	17:43:02	02:43:02	9,20	2,71722222
7	5	15:00:00	17:45:05	02:45:05	9,09	2,75138889
8	6	15:00:00	17:24:05	02:24:05	10,41	2,40138889

фиг. 1.4

За удобство се въвежда допълнителна колона, в която се превръща интервала от време в часове по посочената на фигура фиг. 1.4 формула. След това вече може да се намери и скоростта на всеки участник. Използването на помощни колони и помощни таблици, отделно от основната се използва често за допълнителни изчисления.

D3 =C3-B3

zadachi_time.xls

	A	B	C	D	E	F
1	Компютърна зала		1час=	1,20	лв.	
2	участник	започнал	напуснал	време	цена/лв./	
3	1	17:20	18:30	01:10:00	1,4	1,16666667
4	2	14:25	15:55	01:30:00	1,8	1,5
5	3	16:25	17:12	00:47:00	0,94	0,78333333
6	4	13:40	14:55	01:15:00	1,5	1,25
7	5	15:20	16:00	00:40:00	0,8	0,66666667
8	6	16:15	17:10	00:55:00	1,1	0,91666667

фиг. 1.5

Фигура 1.5 демонстрира решението на четвъртата задача, като превръщането в часове е направено по формулата: F3 = (HOUR(D3)+MINUTE(D3)/60)

5. Особености на употребата на абсолютен адрес във формулите. И при двете предходни задачи се изисква съобразяване на това как да се запишат адресите на клетките, съдържащи пробега и цената за 1 час, във формулите, изчисляващи средната скорост и цената (фиг. 1.4 и фиг. 1.5). Изисква се във формулите– копия да не се променя адреса на клетката, в която се съдържа пробега (цената). Забелязва се, че при използване на относителните адреси на клетките E1 и D1 не изчисляват правилно дължимите суми и скорости. Следователно тези копия, които създава табличният процесор чрез механизма на относителното адресиране не могат да бъдат използвани в конкретната изчислителна ситуация. Решаването на тези и много други таблични задачи изисква копирането на формулите да става без трансформиране на фактическите параметри. Затова фактическият параметър трябва да е абсолютния адрес на клетката, за да остане тя непроменена във всички копия на формулата (съответно E1 за зад. 3 и D1 за зад. 4).

Подобна задача е и следващата, която би могла да се даде за проверка и оценка на знанията, след като са решени и обсъдени в предишни часове сходни примери.

5 зад. Във всяка видеотека се извършва отчетност при отдаване на касети и DVD под наем, чрез документ, в който се отбелязва броя на наетите касети, датите на наемане и връщане от клиента, както и дължимата сума. Наемът за едно денонощие за касета е 1 лв., а за DVD – 2 лв. Да се разработи електронен вариант на отчетния документ.

	A	B	C	D	E	F	G	H	I	J
1		цена на 1 касета	1 лв	цена DVD	2,0 лв					
2	№	Клиент	наема				Дати		Дължи	
3			касети		DVD					
4			брой	номера	брой	номера	наемане	връщане	брой дни	сума
5	1	Иван Петров	1	17	2	11,56	26.3.2006	27.3.2006	1	5 лв
6	2	Мила Пеева	2	15, 23	1	41	26.3.2006	29.3.2006	3	12 лв
7	3	Сия Иванова	4	11, 34, 21, 3			24.3.2006	26.3.2006	2	8 лв
8	4	Иван Тотев	3	7, 58,8	1	52	24.3.2006	27.3.2006	3	15 лв
9	5	Стоян Стоянов	1	89	2	32,12	26.3.2006	28.3.2006	2	10 лв

Фиг.1.6

Друга задача, в която се прилагат функциите IF(…), HOUR(…) и MINUTE(…) е следната:

6 зад. Съставете електронна таблица, която определя продължителността на полетите на една авиокомпания, като се отчитат разликите в часовите пояси, посочени спрямо времето по Гринуич. Часовете на излитане и кацане са в местно време на съответното летище. В колона "Дни" са посочени поредните номера от седмицата, в които се извършва полета. Ако часът на кацане е по-малък от часа на излитане, то полетът е приключил на следващия ден.

Целта е учениците да се научат да правят и по-сложни изчисления в електронна таблица, като ги разделят на поредица от по-лесни стъпки.

K3 =IF(F3<G3;G3-F3+J3;A21+G3-F3+J3)

	A	B	C	D	E	F	G	H	I	J	K
1	от		за		дни	излита	каца	№ на полет	тип самолет	закъснение	Продължителност
2	легище	ч.пояс	летище	ч.пояс							
3	Варна	1	Амстердам	2	1234567	10:55	12:50	KL1883	737		0,0798611
4	Анкара	2	Варна	1	--3-567	14:55	16:40	O2868	M80		0,0729167
5	Атланта	-5	Виена	1	1234567	17:50	09:20	DL12476	340	00:15	0,65625
6	Банкгкок	7	Виена	1	1--4-6-	01:55	06:40	BK61	M11		0,1979167
7	Шанхай	8	Виена	1	--3-5-7	11:00	16:35	CA345	330	00:20	0,2465278

L	M	N	O
1	1	55	0
-1	1	45	2
6	15	45	9
-6	4	45	10
-7	5	55	12

Фиг. 1.7

След като се въведат входните данни трябва да се изчисли продължителността на полета. Това става като от часа на кацане се извади часа на излитане и се прибави закъснението, ако полетът е приключил същия ден. Ако е завършил на следващият ще трябва да прибавим 24 часа. Затова е удачно да се избере една клетка извън таблицата (например A21) и след като се форматира от тип TIME, в нея да се въведе 24:00. Тогава в клетка K3 се въвежда показаната на фигурата формула. От тази стойност трябва да извлечем часа и минутите – в колони M и N, като въведем формулите: в клетка M3 := HOUR(K3) и в клетка N3:=MINUTE(K3). В предната колона L се изчислява часовата разлика: в клетка L3:=D3-B3. Накрая от продължителността на полета в часове трябва да извадим тази часова разлика. В крайна сметка в последните две колони (O и P) се получава колко минути и колко часа е продължил всеки полет. За по-добра прегледност и яснота таблицата трябва да се подреди и форматира допълнително по познати техники.

6. Заключение. За да осъзнаят по-добре необходимостта от обучението по електронни таблици, на учениците трябва да се предлагат интересни познавателни задачи, изискващи обработката на различни типове данни. Затова чрез настоящата работа се прави опит за систематизиране на възловите моменти в овладяването на основните операции с числови

стойности от тип “дата” и “време”, както и адаптиране на научното знание по изграждане на електронни таблици с табличен процесор към конкретните училищни условия.

ЛИТЕРАТУРА

1. **Ангелов** А. – Електронни таблици, Ръководство за решаване на задачи. София,”АСИО”, 1999г.
2. **Холбърг** Б. - MS Excel. София, ”Алекс Софт”, 1997г.
3. **Кънчева** С. , Иванов И. –„Информационни технологии”,сборник със задачи и тестове. София, „Нова Звезда”, 2001г.
4. **Математика и Информатика. Бр. 4** /2000г. – Учебна програма по Информационни технологии за 9 и 10 клас, профилирана подготовка

СОФТУЕРНИ СРЕДСТВА ЗА ОСИГУРЯВАНЕ НА СРЕДА, ПОЗВОЛЯВАЩА ИЗПЪЛНЕНИЕ НА СОФТУЕР ПРЕДНАЗНАЧЕН ЗА РАЗЛИЧНИ ОПЕРАЦИОННИ СИСТЕМИ.

КОНСТАНТИН С. ЦВЕТКОВ, ДЕЛЯН ХР. СЪРМОВ

SOFTWARE SOLUTIONS FOR INSURANCE OF AN ENVIRONMENT, WHICH ALLOWS AN EXECUTION OF SOFTWARE DESIGNED FOR DIFFERENT OPERATING SYSTEMS

KONSTANTIN S. TSCVETKOV, DELIAN HR. SARMOV

The development of the Operating Systems during the last years led to wide variety of software products, designed for different platforms. The incompatibility between the OS usually is a result of the characteristics in their architecture. In the last years have been developed a number of products, which allow an execution of programs of an Operating System, different from the one they are written for.

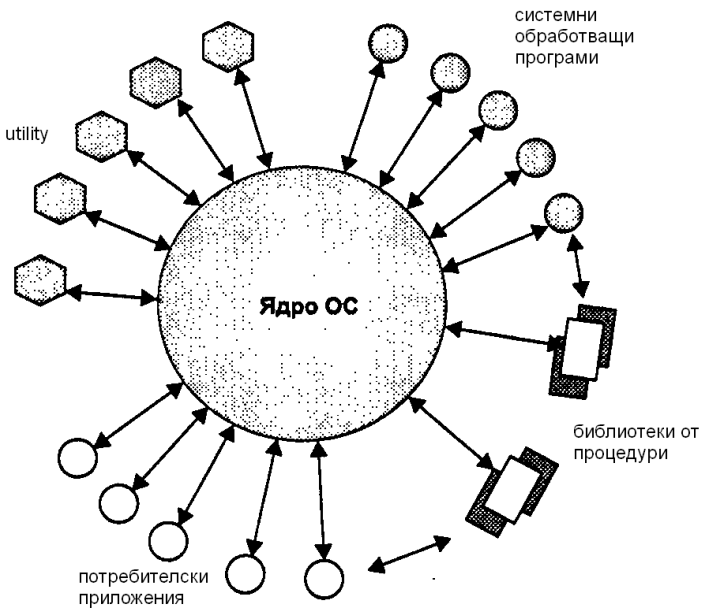
We present the main directions in the architectural principles of the OS. The organization of the nucleus and the different modules has been examined, as well methods, used for a development of an applied environment, in which products are executed, designed for other Operating System.

KEY WORDS: Operating Systems, kernel, Windows, Linux, UNIX, Emulation, Virtual mashine, Exokernel.

Разглеждайки разнообразието на наличния в момента софтуер прави впечатление, че върху създаването на всеки софтуерен продукт работят различни компании, обикновено като конкуренти. По този начин се предлагат алтернативни програми, решаващи определен проблем. Тъй като операционите системи (ОС) също са софтуерен продукт, има разработени редица ОС, които помежду си са несъвместими. В много случаи дори не е налице и пълна обратна съвместимост, тоест със старите операционни системи на същата фирма. В тези условия софтуерът се разработва за определена хардуерна и софтуерна платформа. Потребителят е принуден да избере операционна система, на която да работи и заедно с това той се обвързва да работи само със софтуера предназначен за нея. При изследването на несъвместимостта е важно да се намерят причините поради, които потребителските среди предлагани от ОС са различни.

1. Обща структура на операционните системи.

Съвременните ОС представляват сбор от добре структурирани модулни системи, удобни за развитие, разширяване и пренос на нови платформи [4].



Фиг. 1. Модули на ОС и взаимодействие между тях

Както се вижда на фиг. 1., според предназначението си модулите на операционните системи могат да се разделят на две групи:

1.1. Ядро – модули, изпълняващи основните функции на операционната система;

Модулите на ядрото изпълняват функции като управление на процесите, паметта, вход/изход и други функции, без които функционирането на операционната система не е възможно. Терминът “ядро” се разглежда различно в различните операционни системи. Най-често използвания критерий за да принадлежи един модул на ядрото е да работи в привилегирован режим.

1.2. Модули, изпълняващи спомагателни функции;

Спомагателните модули също се разделят на няколко групи:

1.2.1. Програми решаващи определени задачи по съпровождането на операционната система (utility). Такива са програмите за записване на дискове, архивиране на данни, мониторинг и администриране на системата;

1.2.2. Системни обработващи програми – компилатори, свързващи програми, текстови редактори;

1.2.3. Програми за предоставяне на допълнителни услуги на потребителя. Такива са потребителските интерфейси, калкулатор, както и игри;

1.2.4. Библиотеки от процедури с различно предназначение, облекчаващи разработката на приложения;

За надеждното управление на хода на изпълнението на приложенията операционната система трябва да има определени привилегии в сравнение с приложенията. За тази цел се използват и апаратни средства. Апаратурата на компютъра трябва да поддържа като минимум два режима на работа – потребителски режим (user mode) и режим на ядрото (kernel mode). Тъй като ядрото изпълнява основните функции, тъкмо то най-често работи в привилегирован режим. По този начин именно архитектурата на ядрото и комуникацията на приложенията с него определя архитектурата на операционните системи и е причина за по-голямата част от несъвместимостите между тях.

2. Видове ядра.

2.1. Монолитно ядро;

Монолитният подход дефинира интерфейс от високо ниво върху хардуера, заедно с набор от примитиви и системни извиквания, за да се имплементират системни услуги като мениджмънт на процесите, успоредност и мениджмънт на паметта, в няколко модула, които се изпълняват с правата на супер потребителя (root) [3,4,5,7]. Макар, че всеки модул обслужващ тези операции е отделен като цяло, интеграцията на кода е много тясна и се постига трудно. Обаче, когато имплементацията е пълна и надеждна, тясната интеграция на компонентите позволява особеностите на ниско ниво на основната система да се използват ефективно.

2.2. Микро ядро;

При микроядрата се дефинира проста абстракция над хардуера с набор от примитиви и системни извиквания, за да се имплементират минимални услуги, като мениджмънт на нишки, адресни пространства и комуникация между процеси [4,11]. Главната цел е разделянето на модулите предоставящи услуги от основните компоненти на системата. Например, процеса на заключване на ВХОДА/ИЗХОДА (Input/Output) може да бъде имплементиран в отделен сървърен модул, който ще се изпълнява над ядрото. Сървърните програми, които изпълняват задачи на системата от по-високо ниво, са модулни и опростяват структурата и дизайна на ядрото. Ако някой модул не работи правилно, не се нарушава работата на цялата система. Той може да бъде рестартиран независимо от другите модули. Примери за микро ядра са: AIX, Contiki, Mach, Minix, MorphOS, QNX, RadiOS, BeOS.

2.3. Хибридно ядро;

Хибридните ядра са разработени в резултат на компромис, направен в ранните адаптации на микро ядрата, преди да бъде показано, че микро ядрата също могат да са много бързи [1,2,4]. Хибридните ядра са разширени микро ядра. В тях е включен част от кода, който при ОС с микроядра е извън ядрото. По този начин се постига по-голямо бързодействие, като се запазва модулността характерна за системите с микроядро.

Повечето модерни операционни системи влизат в тази категория. Примери за ОС с хибридни ядра: Microsoft Windows NT, XNU DragonFly, BSD, ReactOS [1,2].

2.4. Външно ядро - Exokernel;

Външните ядра, също известни като вертикално структурирани операционни системи, са нов и радикален подход към дизайна на операционната система [6]. Идеята е разработчикът да направи всички решения, за действието на хардуера. Външните ядра са изключително малки по размер, тъй като тяхната функционалност се ограничава до защита и разпределение на ресурсите.

3. Средства за осигуряване на множествена потребителска среда;

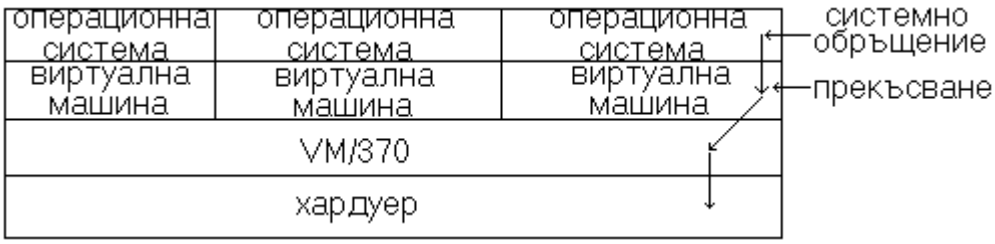
Архитектурните особености на операционните системи касаят само системните програмисти. Поради тази причина за част от потребителите е необходимо да разполагат с достъпни средства за работа с повече от една операционни системи или техни модули. Тези средства могат да се класифицират в четири групи.

3.1. Графичен терминал;

Най-лесният начин за използване на една ОС докато компютърът е под управление на друга е чрез графичен терминал за отдалечен достъп като: Telnet, VNC (който използва RFB (Remote FrameBuffer) протокол за отдалечен достъп до друг компютър) или в някои случаи - X-server. При този вариант е нужно да има втора ОС, инсталирана на втори компютър. Компютърът на потребителя служи единствено за визуализиране на нейния екран и за прехвърляне на клавиатурата и мишката. В този случай всички приложения се изпълняват под пълния контрол на отдалечената ОС.

3.2. Виртуална машина;

Първата виртуална машина е създадена през 1966 г. от изследователска група на IBM в Кеймбридж, щата Масачузетс [8,9].



Фиг. 2. Структура на VM/370

Системата притежава машинно зависима част, която предоставя няколко виртуални машини. Тези виртуални машини не са разширени [8]. Не поддържат файлове и други удобства, а представляват точни копия на апаратурата. Например, режим на ядрото и потребителя, прекъсвания, вход/изход и всичко друго присъстващо в реалната машина. Тъй като всяка виртуална машина е идентична с истинската компютърна система, на всяка може да работи произволна ОС, поддържаща хардуера на този компютър. Когато някоя от операционните системи направи системно обръщение, виртуалната машина го прихваща и изпълнява прекъсване на реалния компютър.

Принципът на работа на виртуалните машини е следния. На реалния компютър се инсталирана ОС наречена host operating system (HOS). На нея е инсталиран софтуер, който управлява виртуалните машини.

Виртуалната машина в повечето случаи представлява образ на файловата система, получаваща се при инсталиране на някаква ОС на избран виртуален компютър. Съхранява се под формата на файл или е разположен в определен дял на твърдия диск. Такава ОС се нарича guest operation system (GOS). Необходимо условие да се инсталира GOS е да са налични драйвери за виртуалния компютър. С помощта на софтуера за управление на виртуалните машини GOS се зарежда в заделено пространство от оперативната памет на реалния компютър. Тя работи с хардуера на виртуалния компютър. Неговите инструкции се прихващат от софтуерът за управление на виртуалните машини и се транслират в код съвместим с реалната апаратура. По този начин виртуалната машина е способна да взаимодейства с апаратурата на компютъра. В повечето случаи е възможно на една компютърна система с виртуална машина на нея да се инсталират няколко различни операционни системи (в общия случай различни), които могат да работят едновременно. За тази цел е необходима достатъчно оперативна памет, която да обезпечи работата на заредените операционни системи.

Съществуват софтуерни продукти (Mac-On-Linux, ShapeShifter, Simics, Parallels Workstation, Plex86, Virtual PC за Windows, и VMware), с чиято помощ може да се създаде един или повече виртуален компютър върху HOS [1,8,10]. На този виртуален компютър може на свой ред да се инсталира друга ОС. С помощта на VMware може при HOS Windows, в прозорец на виртуалната машина да се изпълнява Linux или обратно - основната ОС да е Linux а в прозорец да се изпълнява Windows.

3.3. Емулатор;

При този вариант на операционната система - домакин се изпълнява емулатор, който "лъже" програмите, че се изпълняват под управлението на друга ОС. Най-популярният (и вграден в повечето дистрибуции на Linux) емулатор на Windows под Linux е Wine [7]. Неговата платена надстройка Winex разполага с поддръжка на DirectX. Други емулатори са CoLinux (изпълняващ Linux под Windows), LINE (изпълняващ x86 Linux програми под

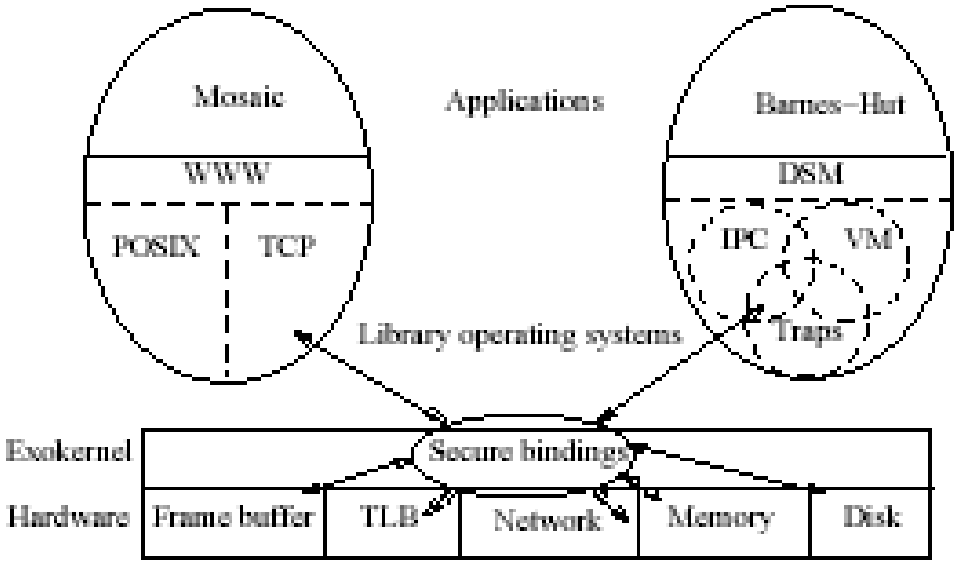
Windows). Някои емулатори като Cygwin (както и други базирани на Cygwin, като MinGW или MinGW32) включват X-server за Windows (Xfree86), както и пакет от инструменти, които позволяват да се прекомпилират приложения от Linux за Windows.

Между Cygwin от една страна и Wine, CoLinux или LINE от друга има съществени разлики. В съответствие с модерната архитектура на операционните системи емулаторите разполагат с библиотеки реализиращи изпълнението на програмите в графичния интерфейс, за който са написани. Тези библиотеки са написани така, че да работят под управлението на HOS. В Cygwin обаче приложенията трябва се прекомпилират за да могат да работят с този емулатор.

Всеки емулатор създава приложна среда, в която да работи друга ОС. Приложната среда е съвкупност от средства предназначени за организация на изпълнението на приложенията използващи определен набор от машинни команди, определен тип графична среда – Application Programming Interface (API) и формат на изпълняваната програма. Всяка ОС създава минимум една приложна среда. При емулацията се решава проблема със съвместната работа на няколко приложни среди в рамките на една ОС. Съвкупността от няколко приложни среди се нарича множествена приложна среда.

Трябва също да се отбележи, че част от операционните системи предлагат вградена поддръжка на други приложни среди, които работят като обикновено приложение [2,4]. Така например в операционната система Windows XP се използва емулация за приложенията написани за MS-DOS. Емулаторът използва виртуален процесор 8086 с 16-разрядна адресация на паметта и ограничение на обема памет до 1 Mb. Системата MS-DOS е затворена в адресното пространство на виртуалната машина 8086, така че емулаторът само препраща прекъсванията към MS-DOS. Когато MS-DOS опита самостоятелно да осъществи входно-изходна операция, тя се прихваща и се изпълнява от емулатора.

3.4. Операционни системи с външно ядро (Exokernel);



Фиг. 3. Примерна система базирана на външно ядро

През 1994 г. се заражда идеята за създаване на малко ядро с ограничена функционалност, което да осигури защита и размяна на ресурси [6]. Външните ядра са значително по опростени от микроядрата и монолитните ядра. Драйверите на устройствата са изнесени извън ядрото. Платформено зависимата част се състои от три групи обекти:

3.4.1. ядро;
3.4.2. драйвери на устройства;
3.4.3. обекти реализиращи механизъм за свързване на обектите.

Платформено зависимата част при тази архитектура включва определен набор от приложни среди и общо пространство за данни. Различните потребителски приложения работят в някоя от приложните среди, използват общото пространство за данни и драйверите на устройствата, а механизма за свързване на обектите осигурява комуникацията между тези компоненти.

Теоретично е възможно под управлението на едно външно ядро да съществуват няколко вида операционни системи (Windows, Linux, UNIX) и разработчикът може да избира да не зачита или да разширява функционалността с оглед на технически характеристики или бързодействие. В момента, външните ядра са предмет предимно на научни изследвания и не се използват в нито една от големите комерсиални операционни системи.

Въпреки високият потенциал на тази идея, тя не намира приложение в големи комерсиални операционни системи и през 2005 г. разработката и е прекратена.

4. Изводи.

4.1. Опитите за решаване на проблема с несъвместимостта все още са свързани предимно с осигуряване на възможност да се използват приложения само на една ОС най-често чрез използване на емулатори.

4.2. Когато целта е да се постигне съвместимост с повече от една ОС, се използват виртуални машини с инсталирани на тях операционни системи. В тази област има комерсиални продукти предлагащи и корпоративни решения.

4.3. Множествените потребителски среди все още се използват рядко, като тяхната работата е скрита в ОС от потребителя.

ЛИТЕРАТУРА

1. **Мейс** Т. Обзор архитектуры Windows 3.x, Windows 95, OS/2 Warp, Windows NT - PC Magazine/RE (C) СК Пресс 1996, №1.
2. **Дэвид** С. Архитектура ядра Windows NT 5.0 – Открытые Системы, 1999, №1
3. **Олифер** В. Г., Х. А. Олифер. Сетевые операционные системы, 2002
4. **Таненбаум** Э.С.,А.С. Вудхалл. Операционные системы Разработка и реализация. Пер. с англ. Д. Шинтяков, 2006.
5. **Maurice** J. B. The design of the UNIX operating system, 1986.
6. **Dawson** R. E, M. F. Kaashoek, J. O'Toole. Exokernel: An Operating System Architecture for Application-Level Resource Management, Cambridge.
7. **Dike** J. A user-mode port of the Linux kernel. In Proceedings of the 2000 Linux Showcase and Conference, October 2000.
8. **Goldberg** R. P. Survey of Virtual Machine Research. IEEE Computer, pages 34-45, June 1974.
9. **Magnusson** P. S., M. Christensson, J. Eskilson, D. Forsgren, G. Hallberg, J. Hogberg, F. Larsson, A. Moestedt, and B. Werner. Simics: A Full System Simulation Platform. IEEE Computer, 35(2):50-58, February 2002.
10. **Sugerman** J., G. Venkitachalam, and B. H. Lim. Virtualizing I/O Devices on VMware Workstation's Hosted Virtual Machine Monitor. In Proceedings of the 2001 USENIX Technical Conference, June 2001.
11. **Hand** S. M. Self-Paging in the Nemesis Operating System. Third Symposium on Operating Systems Design and Implementation, February 22-25, 1999.

ВЛИЯНИЕ СТАБИЛНОСТТА НА СИНХРОНИЗАЦИЯТА ВЪВ ЦВОС ВЪРХУ ВЕРОЯТНОСТТА НА ГРЕШКА

АЛЕКСАНДЪР П. МИЛЕВ

DEPENDENCE BETWEEN JITTER SYNCHRONYZATION AND ERROR RATE

ALEKSANDAR P. MILEV

This paper describes the influence of different factors in digital fiber optic communications over jitter synchronization and error rate and their dependences. It is shown equations and graphics of influence thermal noise, shot noise and pulse pattern over jitter synchronizations. It is evaluated the dependence of synchronization and bit error rate.

KEY WORDS: jitter synchronization, error rate, fiber optic communications

При предаването на сигналите в цифрови системи е необходимо в приемника да се възстановяват не само двоичните елементи, идентични на предаваните, но и информацията за синхронизацията по фаза и честота. Синхронизацията е изключително важна за правилната работа на приемника и за задаване на точното време за определяне на операцията по вземане на решение, цифрово-аналоговото преобразуване и разделянето на информационните канали [1].

Възможно е да се реализира разделно предаване на синхро сигнала. За тази цел е необходимо да се използува друга проводна линия или допълнително оптично влакно. В този случай са нужни две строго идентични трасета за да може да се осигури едновременно пристигане на сигналния импулс и синхронизиращия импулс.

Втория подход, който е и по прост се реализира чрез добавяне на синхросигнал непосредствено към двоичния сигнал. Съществуват множество решения на такъв подход. Едиинствения проблем в този случай е, че след приемане на информационния и синхро сигнала първоначално трябва да се отдели синхрочестотата, на чиято основа се формира момента за вземане на решение за наличие или отсъствие на импулс т.е. за отделяне на 1 и 0 от информационния канал.

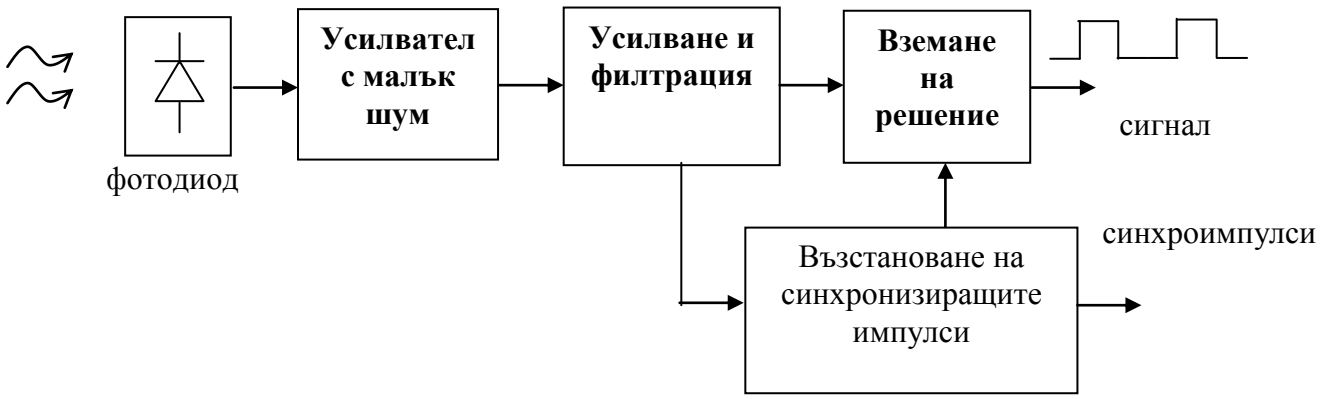
От друга страна изличането на синхро честотата зависи от множество фактори като температурни условия, междусимволна интеференция, т.нар. дрейф на нулата. Всички те изискват контрол по отношение на тяхната стабилност, но поради невъзможност за комплексното им компенсиране , следствие се явява нестабилността на синхронизацията в ЦВОС. Това явление носи наименованието “джитер” и в практиката е необходимо да се съобразяваме с него и да търсим решения за неговото намаляване [2,3].

Става ясно, че от съществено значение е точното извличане на информацията за синхро честотата от поредицата на информационния сигнал.

Основните фактори, влияещи върху стабилността на синхронизацията са:

- неидеалност на тактовите вериги (невъзможност за мигновенна настройка на тактовата тракт, неидеални ограничаващи вериги спрямо термалния и пиковия шум);
- коригираната форма на импулсите (качествения фактор на функцията на корекция, и междусимволната интерференция);
- съвкупността от импулсите на предаваните данни (дължината на думата, броя на идентичните символи и отношението на марката).

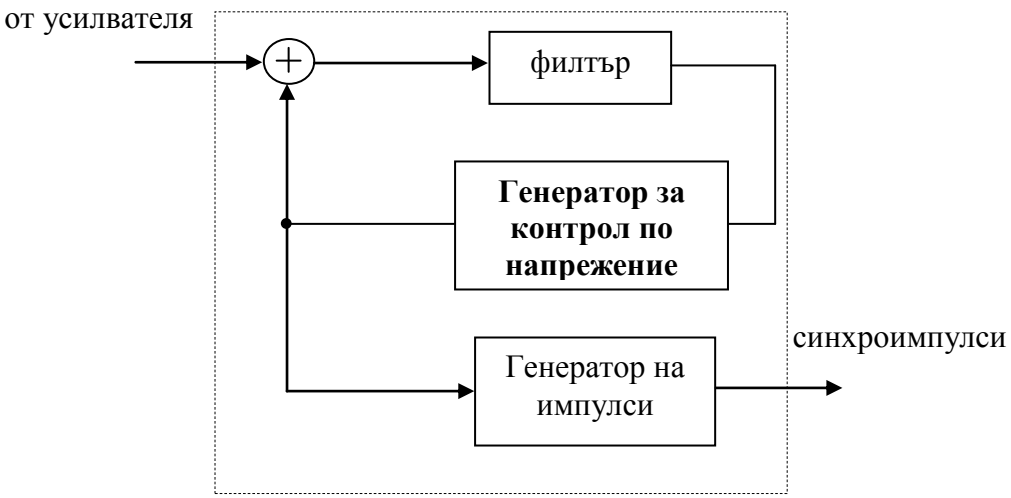
Извличането на синхрочестотата се извършва чрез устройство за възстановяване на на синхронизиращите импулси (фиг.1).



Фиг.1

След отделянето на синхрочестотата се дефинира и момента за вземане на решение за определяне наличието или отсъствието на импулс (1 или 0). Очевидно, е че нестабилността на параметрите на елементите за извличането на синхро честотата би довело и до флуктуацията на момента за вземане на решение. Това от своя страна би довело до нарушение в правилното отчитане на информационните импулси и в резултат ще се повиши вероятността на грешка в системата.

Самата верига за отделяне на тактовата синхро информация може да бъде представена като съвкупност от блокове показани на фиг.2.



Фиг.2

Анализът на фиг.2 показва, че основните показатели, които могат да повлияят на флуктуацията на тактовия интервал са топлинния шум, пиковия шум и последователността от информационни данни (1 и 0). Поради тази причина те трябва да се анализират при отчитане стабилността на синхронизацията и вероятността на грешка.

От съществено значение е и качествения фактор на филтъра, както и предавателната характеристика на филтъра.

За да се отчетат тези фактори ще приемем, че оптичното влакно е достатъчно дълго и че е настъпила междумодова дисперсия и следователно формата на оптическия импулс на входа на приемника е Гаусова

$$m(t) = \frac{\Lambda_e}{\alpha T\sqrt{2\pi}} \exp\left(- \frac{t^2}{2\pi^2 T^2} \right)$$

(1)

където Λ_e е средния брой на електроните за един импулс; αT е разширението на импулса; Т е тактовия интервал.

Предавателната функция на филтъра може да бъде записана във вида

$$Q(f) = \frac{\cosh \beta/T - jT \sinh \beta/T}{1 + j(\pi f/2B_ш)} \exp\left(\frac{\beta f^2 T^2}{2} \right)$$

където $B_ш$ е честотната лента на приемника, $\beta = (2\pi\alpha)^2$

Средният брой на електроните за един импулс лесно може да бъде получен от приетата средна оптична мощност Рср

$$\Lambda_e = \frac{\eta}{h\nu} P_{cp} T$$

където η е квантовата ефективност на фотодетектора (≈ 0.75) и $h\nu$ е енергията на фотона (2.3×10^{-19}J).

Ако с ρ означим отношението на средната мощност на синхронизиращия сигнал и мощността на топлинния шум в рамките на честотната лента (-1/T,1/T) то неговата стойност е свързана с отношението сигнал/шум чрез израза

$$\rho = SNR\left[\frac{\alpha\sqrt{2\pi}}{2} \exp\left(-2\pi^2\alpha^2\right) \right]^2$$

Тогава отчитайки факторите, влияещи върху стабилността на синхронизацията при малки стойности на разширение на импулсите ($\alpha<0.5$) могата да се запишат изрази за флуктуацията на снхронизацията под влиянието на съответните компоненти:

- влияниет на топлинния шум

$$\sigma_t^2 = \frac{B_ш T}{4\pi^2 \rho}$$

- влиянието на пиковия шум

$$\sigma_ш^2 = \frac{B_ш T G_2}{4\pi^2 G_1^2} \left(\frac{2\Lambda_0 / \Lambda_e + 1 - e^{-2\beta}}{\Lambda_e e^{-\beta}} \right)$$

- влиянието на съвкупността от импулси

$$\sigma_и^2 = \frac{B_ш T}{2\pi^2}$$

където Λ_0 е средния брий на първични електрони свързани с тока на тъмно и непълното погасяване на енергията от източника;

G_1 и G_2 са съответно коефициентите на усилване на фотодиода в първоначалния момент и в края на импулса.

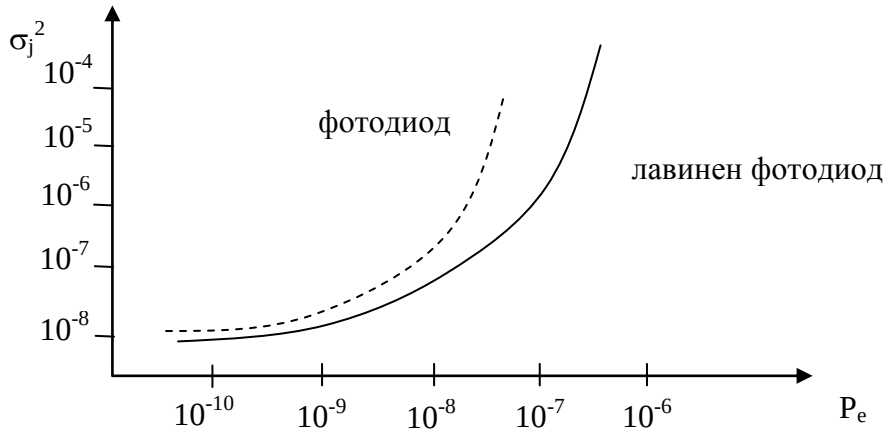
Сумарното въздействие на тези три съставки ще определи общата флуктуация на синхро честотата.

От друга страна можем да отчетем връзката между отношението сигнал/шум и вероятността за грешка чрез израза

$$P_e = \frac{1}{2} erfc\left(\frac{Q}{\sqrt{2}}\right)$$

където Q за цифрови системи е половината от отношението сигнал/шум на входа на приемника.

На фиг.3 е показана зависимостта между стабилността на синхронизацията σ_j^2 и вероятността на грешка P_e.



Фиг.3

Очевидно е, че стабилността на синхронизацията остава в по-големи граници при използуването на лавинен фотодиод, поради по бързата реакция и чувствителността на приемния елемент.

Анализът на изразите за флуктуацията на синхрочестотата показва, че тя нараства с включването на повече повторители, т.е. с увеличаването на броя на регенерационните участъци, а съответно и общата дължина на връзката.

За намаляване или подтискане на флуктуацията може да се използува скремблиран сигнал или сигнал с отношение на марката близко до 0,5.

ЛИТЕРАТУРА

1. **Гауэр** Дж., Оптические системы связи. Москва, “Радио и связь”, 1989.
2. **Козанне** А., Оптика и связь. Москва , ”Мир” 1984.
3. **Швецов** Э. А., Белкин М.Е., Фотоприемные устройства волоконно-оптических систем передачи. Москва, “Радио и связь”, 1992.

УПЛЪТНЯВАНЕ НА ИНФОРМАЦИОННИЯ КАНАЛ В ОПТИЧНИ СРЕДИ ЧРЕЗ МНОГОПОЗИЦИОНЕН СИГНАЛ С ТРИ НИВА

АЛЕКСАНДЪР П. МИЛЕВ

ABOUT USING MULTIPOSITION THREE LEVEL SIGNAL AT FIBER OPTIC SYSTEM FOR INCREASING INFORMATION CAPACITY

ALEKSANDAR P. MILEV

This paper describes an approach of using multilevel signal combined with multiple positions of the third level. It is shown that information capacity has been increased due to including additional parameter for another informational canal.

KEY WORDS: multipostion signal, multilevel signal, fiber optic communications

В цифровите оптични комуникации по многомодово оптично влакно разширението на импулса става проблем при скорости на предаване над 100 Mbps. Дали импулсите ще се изравняват или не междусимволната интерференция в следствие на изкривяванията на влакновите закъснения значително намаляват дължината на регенерационния участък [1].

Известно [2] е, че съществува граница за скоростта на предаване и дължината на оптичното влакно, която не може да бъде превишена без да се отрази на сложността на изравнителя и наличната оптична мощност.

За да преодолее тази граница и за да се увеличи скоростта на предаване можем да използуваме многонивови сигнали.

Скоростта за предаване на информацията за M-нивов сигнал, ако всички нива са еднакво равновероятни е $R=\log_2(M/T)$, където T е цифровия интервал.

Многонивовите сигнали могат да бъдат считани като средство за задоволяване на различни изисквания по отношение кода за предаване (напр. откриване на грешка, липса на НЧ компненти и др.) в прости и ефективни случай дори когато регенерационния участък е ограничен поради високи загуби на енергията. Съществуват много случаи на приложение на сигнал с три нива, които са доказали свойте предимства.

Статията разглежда един подход за приложение на тринивов сигнал, чрез който се реализира многопозиционно кодиране.

При многопозиционните сигнали с относително модулиране информативен параметър се явява броя на интервалите между изпращането на импулс. Тъй като при относителна модулация се съкращава времето за предаване на информация, то за формиране на многопозиционен сигнал с изравняване по време се изпълнява преместване на отделни импулси. Обиковено за изравняването по време, всеки следващ сигнал, оказващ се в пределите на L -позиционен интервал, заедно с предходния сигнал се премества на указания брой m позиции.

Интензивността на информационния многопозиционен сигнал при условие на гаусова форма на импулса може да се представи във вида [3]

$$\lambda(t)=\sum_{j=1}^{N} Ae^{-\alpha\left[t+\tau/2-\left(k_j+(j-1)L\right)\tau\right]^2} \qquad (1)$$

където A – енергията на импулса; τ - продължителност на импулса; j - номер на предавания сигнал, k_j – номер на такта от отчетната точка; L – брой позиции на сигнала; α - параметър отчитащ разширението на сигнала.

В случай на m премествания неговата интензивност може да бъде представена във вида

$$\lambda(t) = \sum_{j=1}^{i_1-1} Ae^{-\alpha\left[t-\tau/2-\tau\sum_{k=1}^{j} n_k\right]^2} + \sum_{j=i_1}^{i_2-1} Ae^{-\alpha\left[t-\tau/2-L\tau-\tau\sum_{k=1}^{j} n_k\right]^2} + \ldots + \sum_{j=i_m}^{N-1} Ae^{-\alpha\left[t-\tau/2-mL\tau-\tau\sum_{k=1}^{j} n_k\right]^2} \quad (2)$$

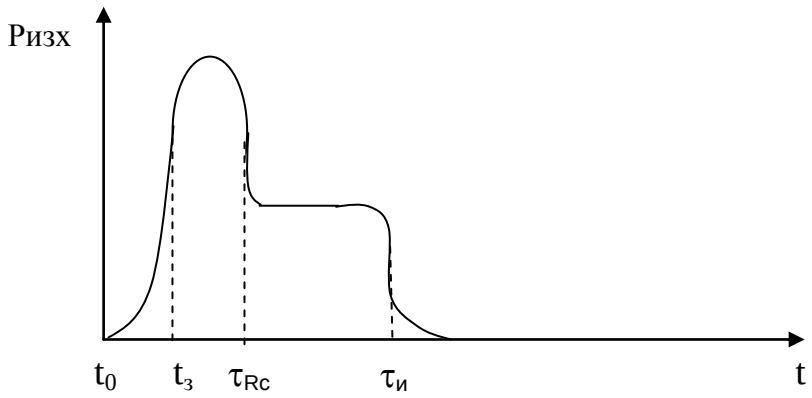
където n_k - броя на позициите между импулсите; i_m – броя на преместваните импулси.

В израз (2) значението на k_j се изменя в интервала 1 – L и се явява номер на такта от отчетната точка. При това k_j се определя като част от случайна величина n_j , която се явява независима за всяко n_k . Затова k_j също е независимо. Следователно за k_j може да се извърши усредняване.

Известно е, че при оптическите системи е необходимо да се въвежда праг за вземане на решение за наличие или отсъствие на импулс (1 или 0) с цел защита от определени шумове (топлинен, от тока на тъмно, междусимволна интерференция). Величината на прага най вече е свързана с флуктуация на интензивността на междусимволните смущения. Праговото ниво се намалява с увеличение разширението на импулса и съществено може да ограничи възможностите на приемника. Очевидно е и че определено отношение сигнал/шум на входа на приемника и наличието на прагово ниво ще бъдат в зависимост от лентата на пропускане на канала и стремежа за търсене на оптимално решение между тях, ще зависи изцяло от вида на сигнала.

Комбинирайки предимствата на тринивовия сигнал и многопозиционната модулация може да се разгледа случай, при който третото ниво от сигнала се модулира относително предходното подобно ниво чрез определен брой тактови интервали m.

Нека представим тринивовия сигнал с форма показана на фиг.1.



Фиг.1

Импулсът е задържан на време $t_з$, определено от израза [3].

$$t_з = t_0 \cdot \ln\left(\frac{I_{pRc}}{I_{pRc} - I_{пр}}\right) \quad (3)$$

където t_0 – началният момент на включване.

τ_{Rc} - продължителност на импулса Rc;

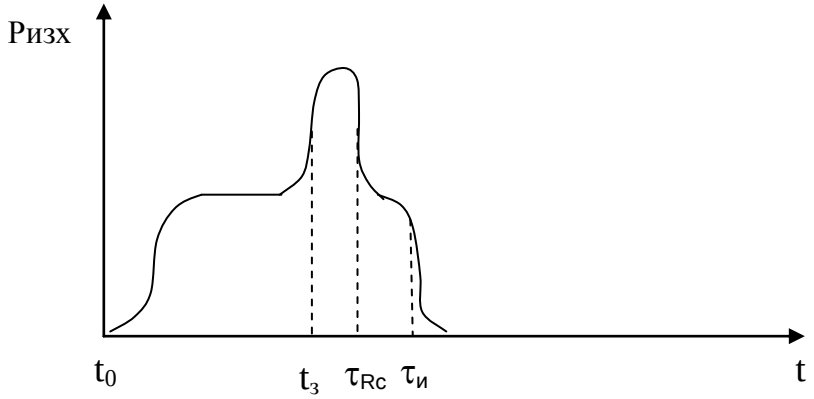
$\tau_и$ - продължителност на импулса за данни за ниво две;

I_{pRc} - тока, необходим за формиране на импулса за тиво три;

$I_{рд}$ - тока , необходим за формиране на един бит информация.

Имайки предвид възможната реализация на сигнал с показаната на фиг.1 форма то може да направим извода, че импулса Rc може да бъде преместван на различни позиции с цел вмъкване на допълнителна информация. Местоположението на този импулс Rc се явява носител на информация, в зависимост от това какво е неговото отместване по отношение на от предходния подобен импулс.

Импулсът Rc може да бъде поставен както в началото на импулса за ниво две, така и в неговия край, както е показано на фиг. 2.



Фиг.2

Нека дефинираме два съседни импулса Rc1 и Rc2 единият, от които е в началото, а другия съответно в края на импулса с ниво две $\tau_и$. Нека разположим тези два импулса на два съседни тактови импулса с ниво две $\tau_и$ По такъв начин може да се твърди ,че позицията на импулса Rc спрямо импулса за данни с ниво две $\tau_и$ ще бъде носител на информация съглано таблица 1.

Табл.1

Местоположение на двойка съседни Rc импулса	Rc1	Rc2
0 0	-	-
0 1	-	+
1 0	+	-
1 1	+	+

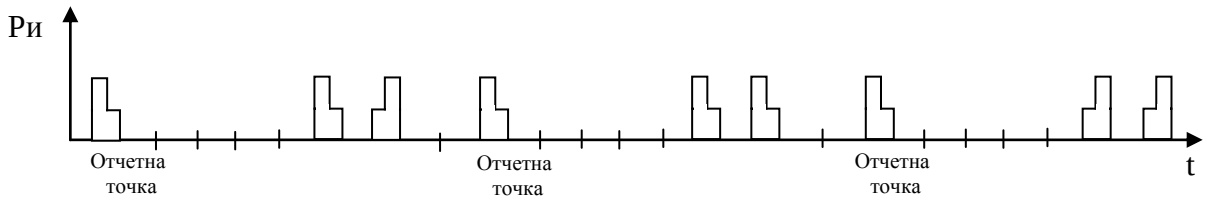
Възможната реализация е за 4 позиционен тринивов код. В този случай групата импулси Rc1 и Rc2 отстоят винаги на едно и също разстояние (еднакъв брой тактове). Това от една страна е добре, защото този факт би подпомогнал за подобряване точността на синхронизацията в комуникационните системи тъй като открояването на двойката импулси ще бъде индикатор за самосинхронизиране и по точно сфазиране на отделящите синхрочестотата генератори.

Отчитайки възможното различно отместване на импулса Rc спрямо предходния подобен импулс Rc, то може да се осигури още един параметър като носител на информация,

а именно броя на тактовете, на които са отместени двойката импулси Rc1 и Rc2 спрямо определена отчетна точка. При този подход може да дефинираме повече от 4 позиционен код с три нива.

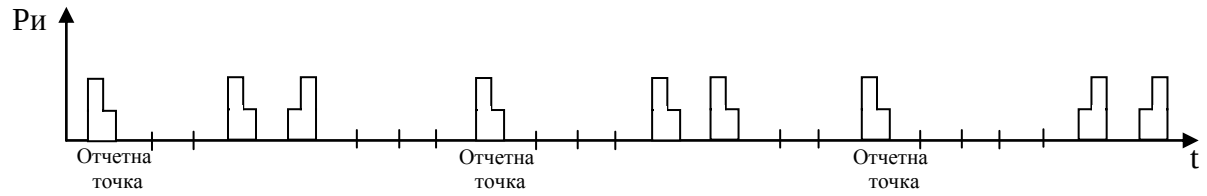
Нека да дефинираме отчетна точка като импулс Rc , който се явява в началото на всеки кадър информация (или в началото на блоков код, примерно mBnB). Тогава броя на тактовете, на който отстоят импулсите Rc1 и Rc2 спрямо отчетната точка се явява носител на допълнителна информация.

При еднакво отстояване на двойката импулси Rc1 и Rc2 съвкупността от информационни сигнали ще бъде като поазаната на фиг.3.



Фиг.3.

При различно отместване на имулсите Rc1 и Rc2 информационните сигнали ще бъдат подобно на тези от фиг.4.



Фиг. 4.

В зависимост от степента на отместване от стандартна точка (такт) спрямо първоначално избраната отчетна точка (такт, начало на блочна кодова комбинация mBnB) и формирания брой позиции, може да бъдат разгледани 8, 16 и 32 позиционни кодове, (според стойността на m и n).

С осигуряването на повече информация за предване за единица време (дължината на един блоков код mBnB) се гарантира повишаване на информационната способност на влакново оптичните комуникационни среди.

Допълнително предимство от използуването на тринивов сигнал е увеличаването на разстоянието между регенераторите на сигнала заедно с увеличаването на скоростта на информацията, следователно и далечината на реализуемата комуникация.

ЛИТЕРАТУРА

1. **Гауэр** Дж., Оптические системы связи. Москва ,"Радио и связь", 1989.
2. **Техника оптической связи**: Фотоприемники / Под ред. У. Тсанг, Москва ,"Мир", 1988.
3. **Швецов** Э. А., Белкин М.Е., Фотоприемные устройства волоконно-оптических систем передачи, Москва, "Радио и связь", 1992.

МОДЕЛИРАНЕ НА ИНФЛАЦИЯТА ВЪВ ФИНАНСИТЕ

КРЪСТЮ АНГЕЛОВ КРЪСТЕВ

MODELLING INFLATION IN FINANCE

KRASTYU ANGELOV KRASTEV

This paper is dedicated to a phenomenon in the economies – the inflation. It presents different theories and approaches, which try to explain this occurrence and offer possible ways for modelling it in order to achieve better understanding. Staritng point are the theories of Keynes, Fisher and Pigou and their successors. Later on the paper addresses some aspects of inflation in panel data. In a further step the article shows the latest development of the trials to model inflation. At this stage the focus is on the works of Jarrow-Yildirim, Mercurio and Belgrade, Benhamou and Koehler.

KEY WORDS: Inflation, Modelling, Stochastic Process, Real Interest Rate, Forward Interest Rates, Martingale, Exchange Rate, Risk, Equilibrium

1. Увод

Изходна точка при моделирането на инфлацията във финансовата наука са макроикономически модели, чиято цел е да отразяват и анализират промяната в ценовото равнище и/или индекса на потребителските цени. Литературата по въпроса е обемна, но в нея могат да бъдат откроени две направления:

- изследвания, които се занимават с измерване и обяснение на инфлацията и
- изследвания, чиято цел е да предложат насоки за провеждане на монетарната политика и по такъв начин да дадат своя принос за контролиране и управление на дългосрочната промяна в ценовото равнище.

Първата категория модели си служи с експлицитни макроикономически променливи. При тези подходи се приема, че инфлацията е породена от различни ефекти, като този на постоянно увеличаващото се търсене например, което от своя страна води до нарастване на ценовото равнище.

Настоящата статия си поставя за цел представянето на основните подходи в литературата при моделиране на инфлацията. В началото са изложени класическите модели на Fisher, Keynes и Pigou, които служат като изходна база за по-късните опити за анализ на промяната в ценовото равнище. Като следваща стъпка е обърнато внимание на инфлацита в рамките на времевите редове с произтичащите от това особености при моделирането и е представено развитието на горните подходи в рамките на неокейнсианството и монетаризма. В последната част на статията са изложени модерни подходи, като тези на Jarrow-Yildirim, Mercurio и този на Belgrade-Benhamou-Koehler, които за момента оказват значително влияние върху литературата и оставят отпечатък върху изследванията по въпроса.

2. Ранни опити за анализ на инфлацията

Сред подходите във втората група, следва да се спомене този на Fisher[34]. Моделът дефинира паричното предлагане за един период като кумулативната цена на всички продадени през него стоки и услуги. Този поглед може да се разглежда в светлината на

[34] Fisher, I. The Purchasing Power of Money: Its Determination and Relation to Credit Interest and Crisis. New York 2006, p. 14-27.

инфлация-търсене, която е анализирана от монетаристите[35] посредством допълнителното генериране на парични средства.

Keynes[36] и Pigou[37] предлагат подобно обяснение на проблема. Те представят паричното предлагане M_t като произведение от реципрочната стойност на скоростта на парите в обръщение k_t, равнището на цените в икономиката P_t и националния доход Y_t:

$$M_t = k_t P_t Y_t.$$

Този модел отчита дългосрочната динамика на инфлацията. Последната може да бъде дефинирана като разлика между прираста на паричната маса и растежа на брутния вътрешен продукт. Уравнението на Fisher[38] използва тази динамика и описва инфлацията като разлика между номиналния $\left(r^n\right)$ и реалния лихвен процент $\left(r^r\right)$:

$$i = r^n - r^r.$$

Горните модели не са подходящи за задълбочен анализ, тъй като не игнорират възможностите за арбитраж и не могат да бъдат напаснати към форуърдните реализации на индекса на потребителските цени. Същевременно могат да бъдат отбелязани изследвания, които стъпват на горните модели и водят до добри резултати при изследване на инфлацията в бъдеще. Те могат да бъдат използвани за компенсиране на липсващи пазарни данни, като например за форуърдни значения на краткия край на кривата. Тези модели се характеризират със стационарност и отразяват сезонните колебания на реда. Последните са често входящи данни за финансовите модели. Предимството на тези подходи се състои в това, че те описват структури на зависимост и волатилност на информацията, които не могат да бъдат наблюдавани непосредствено на пазара. Те използват експлицитни променливи, като някои от най-новите проявления на инфлацията. По тази причина тези модели представляват също така първи опит за преминаване към динамично моделиране, въпреки че трудно успяват да комбинират краткосрочната и дългосрочната инфлация.

3. Инфлация във времевите редове

Често използван подход във финансовата математика е непрекъснатото моделиране във времето. Трудност при него представлява преминаването от дискретни променливи към непрекъснати. Проблемът се свежда до намиране първоначално на модел, който да отчита значенията на индекса на потребителските цени в дискретни точки от времето. На тази база следва да се дефинира и използва еквивалентен модел, отразяващ непрекъснатата природа на явлението (стохастично разпръскване) при оценка на индексирани деривати, например. Въпросът относно стационарността на инфлацията е решаващ при избора на модел. При наличието на нестационарен ред се налага диференцирането му минимум два пъти, така че да се гарантира стационарност.[39]

Като алтернатива на горните модели може да се посочи еднофакторният модел на Vasicek[40]. Той описва движението на лихвените проценти, които се изменят в зависимост от пазарните условия. Според модела лихвените проценти следват следната динамика, описана от диференциално уравнение:

---

[35] Friedman, M. The Quantity Theory of Money: A Restatement. B: Studies in the Quantity Theory of Money. Chicago 1956, p. 3-21.

[36] Keynes, J. The General Theory of Employment, Interest and Money, Reprint 1997, p. 292-312.

[37] Pigou, A. The Economics of Welfare. New York 2005, p. 44-62.

[38] Fisher, I. The Purchasing Power of Money: Its Determination and Relation to Credit Interest and Crisis. New York 2006, p. 14-27.

[39] Juselius, K. Domestic and Foreign Effects on Prices and the Efficiency of Single-Equation Analysis. Journal of Econometrics 52/1992, p. 389-402.

[40] Vasicek, O. An Equilibrium Characterisation of the Term Structure. Journal of Financial Economics 5/1977, p. 177-188.

$$dr_t = a(b - r_t)dt + \sigma dW_t.$$

Тук W_t е процес на Wiener, моделиращ пазарния рисков фактор, а е стандартното отклонение, а σ отразява волатилността на лихвения процент. Дрифтът $a(b - r_t)$ показва очакваната промяна на лихвените проценти във време t. Параметърът b е онзи лихвен процент, към който настоящият се стреми в дългосрочен план, така че да се намира в равновесие. При липсата на шокове $(dW_t = 0)$ лихвеният процент остава константен $b = r_t$. Параметърът a трябва да бъде стабилна величина, тъй като той регулира скоростта на адаптация.

Подобно на облигациите с нулев купон, които могат да бъдат напълно определени посредством един-единствен лихвен процент, ценовата динамика в икономиката може да се опише чрез моментната реализация на индекса на потребителските цени (CPI):

$$\frac{dCPI}{CPI} = i_t dt.$$

Подобен подход може да се използва и при анализа на потоците от индексираните облигации например. Купоните на тези ценни книжа се описват със следната формула:

$$c_k = c\left[1 + i_k\left(T_k - T_0\right)\right] = c\frac{CPI\left(T_k\right)}{CPI\left(T_0\right)},$$

където с е купоновият лихвен процент, а i_k е инфлацията за периода $[T_0, T_k]$. Последната е дефинирана като:

$$i_k \Delta T = \frac{\Delta CPI}{CPI} = \frac{CPI_k - CPI_0}{CPI_0} \approx \ln\left(CPI_k\right) - \ln\left(CPI_0\right).$$

За съжаление i_k не може да бъде моделирано непосредствено. В случая се налага да се направи избор между дискретните значения на инфлацията и краткосрочната инфлация $(i_t)_{t\geq 0}$, която не е стационарен процес. Това означава, че последната ще се променя значително, дори и в много малки интервали от време. Този подход има предимството, че може да гарантира дългосрочно равновесие между променливите.

Този модел на инфлацията се използва предимно в банковата индустрия при оценката на суапи и опции върху инфлацията.[41] Допускането за нормално разпределение на реализациите на инфлацията и детерминистичната волатилност правят възможно прилагането на формулата на Black-Scholes за оценка на опции. В случая се налага оценка на параметрите на модела, което не е лесна задача, дори и ако се използва корелацията между различните значения на индекса на потребителските цени.

Често в литературата се срещат модели за инфлацията, които се основават на ограничена информация. Примери за такива модели са нео-кейнсианската крива на Филипс и монетарни подходи към проблема.

4. По-нови опити за обяснение на инфлацията

4.1. Нео-кейнсианска крива на Филипс

Нео-кейнсианската крива на Филипс обяснява инфлацията през настоящия период с очакваната промяна на ценовото равнище през следващия период $E\left(\Delta p_{t+1} \middle| \mathrm{I}_t\right)$ и пределните разходи x_t :

$$\Delta p_t = b_{p_1} E\left(\Delta p_{t+1} \middle| \mathrm{I}_t\right) + b_{p_2} x_t.$$

[41] Koehler, E. and N. Belgrade. Evaluation de produits liés à l'inflation. Paris 2002; Beletski, T. Mathematical Models for Inflation and Pricing of Macro Derivatives. Kaiserslautern 2004.

Основавайки се на подходите на Calvo[42] и Galì и Gertler[43] се стига до вариант на горното уравнение, според който променливата x_t представлява пределните реални разходи за фирмите. Подобни модели допускат скокове на инфлацията. Това води до появата на хибридни нео-кейнсиански криви на Филипс, които предполагат, че икономическите агенти гледат към бъдещето, но и не забравят миналото. Galì и Gertler модифицират горното уравнение и стигат до:

$$\Delta p_t = b_{p_1}^f E\left(\Delta p_{t+1} | \mathrm{I}_t\right) + b_{p_1}^b \Delta p_{t-1} + b_{p_2} x_t .$$

Нео-кейнсианската крива на Филипс се появява като резултат от нарастващите атаки от теоретично и емпирично естество. Bårdsen et al.[44] показват, че кривата зависи и от спецификацията на процеса x_t. Те демонстрират, че е полезно използване на система от уравнения при оценката, вместо едно единствено уравнение. Rudd и Whelan[45] показват, че тестовете, които служат като основа за твърденията на Galì и Gertler, не са достатъчно убедителни и че не е възможно да се направи разлика между поглед напред и назад от страна на икономическите агенти по отношение на инфлацията. Rudd и Whelan разработват нови тестове, с помощта на които успяват да елиминират част от недостатъците на съществуващите подходи.

4.2. Монетарни модели

В P*-модела[46] дългосрочното равновесие на цените резултира от парите в обръщение m_t, в случай че произведеното количество стоки и услуги в икономиката съответства на това при пълна заетост y_t^* и скоростта на парите $v_t \equiv p_t + y_t - m_t$ се намира в равновесие v_t^*: $p_t^* \equiv m_t + v_t^* - y_t^*$. На тази база моделът има следния вид:

$$\Delta p_t = E\left(\Delta p_t | \mathrm{I}_{t-1}\right) + \alpha_p \left(p_{t-1} - p_{t-1}^*\right) + \beta_z z_t + \varepsilon_t ,$$

където основните фактори са очакванията за инфлацията $E\left(\Delta p_t | \mathrm{I}_{t-1}\right)$, разликата между наблюдаваното ценово равнище и това при пълна заетост в икономиката $\left(p_{t-1} - p_t^*\right)$ и други фактори z_t.[47]

С цел изчисляване на разликата в ценовото равнище при пълна заетост и наблюдаваното в момента се налага апроксимацията на произведеното равновесно количество стоки и услуги в икономиката y_t^* и на равновесната скорост на парите в обръщение v_t^*. Ценовата разлика може да бъде представена като разлика между скоростта и произведеното количество стоки и услуги:

$$\left(p_t - p_t^*\right) = \left(v_t - v_t^*\right) - \left(y_t - y_t^*\right).$$

[42] Calvo, G. Staggered Prices in a Utility Maximizing Framework. Journal of Monetary Economics 12/1983, p. 383-398.

[43] Galì, J. and M. Gertler. Inflation Dynamics: A Structural Econometric Analysis. Journal of Monetary Economics 44/1999, p. 233-258.

[44] Bårdsen, G., E. Jansen and R. Nymoen. Econometric Evaluation of the New Keynesian Phillips Curve. Oxford Bulletin of Economics and Statistics 2004, p. 27-46.

[45] Rudd, J. and K. Whelan. New Tests of the New Keynesian Phillips Curve. FRB Working Paper 2004.

[46] Hallman, J., R. Porter and D. Small: Is the Price Level Tied to the M2 Monetary Aggregate in the Long Run? American Economic Review 81/1991, p. 841-858.

[47] Eitrheim, Ø. Testing the Role of Money in the Inflation Process. Norges Bank 2003.

5. Съвременни обяснения

5.1. Модел на Jarrow-Yildirim

Моделът на Jarrow-Yildirim е един от традиционните и най-известни опити за моделиране на инфлацията. Той допуска, че индексът на потребителските цени е разменен курс между номиналните и реални доходности. Това предполага трифакторен модел, основните фактори в който са номиналните, реалните доходности и инфлацията. Приносът на Jarrow-Yildirim се състои в елиминирането на възможностите за арбитраж между факторите в съответствие с условията на модела на Heath-Jarrow-Morton (HJM). Той представлява възможност за анализ на развитието на форуърдните лихвени проценти. По такъв начин се достига до:

$$\begin{cases} \dfrac{dCPI_t}{CPI_t} = \left(r_t^n - r_t^r\right)dt + \sigma_t^{CPI} dW_t^{CPI} \\ dr_t^k = a^k\left(b^k - r_t^k\right)dt + \sigma_t^k dW_t^k,\ k \in \{n, r\}. \end{cases}$$

В това уравнение r^n и r^r са съответно номиналният и реалният лихвен процент с факторни корелационни коефициенти $d\left\langle W_t^k, W_t^j \right\rangle = \rho_{k,j} dt,\ k,\ j \in \{n, r, CPI\}.$ При разглеждането на този модел следва да се отчита, че дрифтът на променливите може да бъде изразен като функция на волатилността.

Отчитайки динамиката на капиталовите пазари, могат да бъдат отбелязани и други фактори, които водят до развитие на оценъчни модели. Инвеститорите са готови да платят значителни суми, за да получат суапи, които да отговарят напълно на нуждите им. Липсата на платци на инфлация например кара инвестиционните банки да плащат инфлация с помощта на инфлационни суапи и да понасят сами риска от несъвпадението на срокове и други параметри Това обяснява, защо равновесната инфлация е по-висока на пазара за суапи отколкото на този за дългови инструменти.

Проблем при модела на Jarrow-Yildirim е възможността за наблюдаване на негативни номинални лихвени проценти. Това води до търсенето на нови модели.

5.2. Модел на Mercurio

Моделът на Mercurio[48] се основава на този на Jarrow-Yildirim. В него могат да бъдат открити два подхода за подобряване резултатите от моделирането на инфлацията и на номиналните лихвени проценти.

Mercurio запазва аналогията с разменното съотношение и разглежда номиналните и реалните форуърдни лихвени проценти, като ги свързва посредством облигации с нулев купон. Това се дължи на възможността за извеждане на модела от анализа на Heath-Jarrow-Morton. Mercurio критикува подхода на Jarrow-Yildirim, като застъпва тезата за наличие на негативни лихвени проценти.

Той предполага log-нормална динамика на номиналните и реалните форуърдни лихвени проценти:

$$\begin{cases} dF^k(t, T_{i-1}, T_i) / F^k(t, T_{i-1}, T_i) = \sigma_{k,i} dZ_i^k(t) \\ B^k(t, T_i) / B^k(t, T_{i-1}) = 1/\left(1 + \tau_i F^k(t, T_{i-1}, T_i)\right),\ k \in \{n, r\}. \end{cases}$$

[48] Mercurio, F. Pricing Inflation-Indexed Derivatives. Working Paper, Milano 2004.

В тези уравнения $F^n(t, T_{i-1}, T_i)$ и респективно $F^r(t, T_{i-1}, T_i)$ са значенията на форуърдните лихвени проценти между T_{i-1} и T_i, $Z_i^n(t)$ и $Z_i^r(t)$ са Браунови движения, отчитащи вероятността за номинално, респ. реално форуърдно плащане. $\tau_i = T_i - T_{i-1}$ и $B^k(t, T)$ представляват стойността във време t на номинална (k=n), респ. реална (k=r) облигация с падеж Т.

Mercurio прави възможно по-доброто разбиране на параметрите на модела и по-удачен избор на данни за верифицирането му. Той използва резултат от подхода на Jarrow-Yildirim: очакваната стойност на форуърдното инфлационно съотношение може да бъде изразена посредством реална и номинална облигация с нулев купон. По тази причина се налага дефинирането на два процеса с мартингали, които могат да бъдат изразени с помощта на номинални и реални форуърдни лихвени проценти. В тях се използват волатилността на процесите и корелационните коефициенти, които са имплицитни функции на първоначалните параметри на модела на Jarrow-Yildirim.

За поток с падеж във време T_n при първия мартингал се получава израз, който е аналогичен на форуърдното разменно съотношение по подобие на подхода, правещ аналогията с разменен курс:

$$I_t(t) = CPI_t \frac{B^r(t, T_i)}{B^n(t, T_i)}.$$

В това уравнение $B^n(t, T_i)$ и $B^r(t, T_i)$ са реалната и номиналната стойност на облигацията с нулев купон във време t, като тя е със срок T_i. Значението на символите тук е същото както горе.

При втория подход процесът се анализира при друга вероятност на бъдещата дата. Тази промяна налага корекцията му и респективно промяна в очакваната форуърдна инфлация. Корекцията от своя страна е функция моментната волатилност на номиналния лихвен процент и на индекса на потребителските цени и на корелацията между номиналните форуърдни лихвени проценти.

При този подход номиналните лихвени проценти винаги са положителни величини, а използваните при моделирането променливи са непосредствено свързани с инфлацията на пазара за деривативни инструменти. Важно предимство на модела на Mercurio се състои в това, че се използват по-малко параметри в сравнение с модела на Jarrow-Yildirim.

5.3. Модел на Belgrade-Benhamou-Koehler

Търгуваната форуърдна инфлация на пазара се различава от тази въз основа на иконометрични прогнози на базата на статистически редове. Belgrade-Benhamou-Koehler предлагат свързването на облигациите с нулев купон и инфлационните деривати с помощта на пазарен модел. Те избират да моделират форуърдните стойности на индекса на потребителските цени. Тези процеси могат да бъдат разглеждани като мартингали при определени вероятности. Авторите допускат, че всяко фиксирано значение на индекса във време T_i, което може да се наблюдава във време t е пределен мартингал

$$\frac{dCPI(t, T_i)}{CPI(t, T_i)} = \sigma(t, T_i) dW_{T_i}^i(t)$$

с детерминистична волатилност $\sigma(t, T)$. Брауновите движения $\{W_{T_i}(t), i = 1, \ldots, n\}$ са корелирани, като връзката може да се представи по следния начин:

$$d\left\langle W_t^i, W_t^j \right\rangle = \rho_{i,j}^{INF}\, dt, \text{ където } i, j = 1, \ldots, n.$$

Доходностите от номиналните облигации с нулев купон са log-нормално разпределени с детерминистична волатилност. Последната е свързана с волатилността на форуърдния индекс на потребителските цени посредством променящата се във времето структура:

$$d\left\langle W_t^i, W_t^j \right\rangle = \rho_i^{INF,\,DF}\, dt.$$

За да включат сезонността в анализа, авторите предлагат два метода за оценка и отразяване на сезонния компонент, който може да се раздели на стохастична и нестационарна част.[49] Според тях се налага корекция на форуърдната стойност на индекса на потребителските цени на дата D_T от година T, като се интерполира сезонността. Предполага се, че индексът на потребителските цени е функция f на скрити компоненти: детерминистичен тренд Z, сезонност S и остатъчен член ε:

$$CPI(t, t) = f(Z_t, S_t, \varepsilon_t).$$

Пазарният модел има редица предимства. Той моделира директно стоящите в основата променливи: волатилността на облигациите с нулев купон и корелациите между значенията на индекса на потребителските цени в две точки от времето. Моделът прави връзка между лихвените проценти по облигациите с нулев купон и тези, които носят доходи през всеки период. Изборът на хомогенна форма на волатилността позволява намаляване на степените на свобода на модела. Той включва и сезонните фактори при напасване на входящите величини. Многодимензионалната характеристика на модела го прави сложен за прилагане и резултатите се получават бавно. Без достатъчно добра параметризация на волатилността, той има прекалено много степени на свобода. По тази причина се налага отчитането на параметричната волатилност.

6. Заключение

В исторически план първите модели за инфлацията се основават на количествената икономическа теория при обясняването на инфлацията или на динамиката на индекса на потребителските цени. Тук могат да бъдат отбелязани обясненията на Keynes, Pigou и Fisher от една страна и тези на по-късните монетаристи и нео-кейнсианци. След появата на пазарни данни моделите се усъвършенстват. Първият модел на Jarrow-Yildirim от 2000 г. изследва инфлацията посредством номиналните и реалните лихвени проценти. Той обаче се основава на ненаблюдаеми параметри и исторически оценки.

Mercurio представя два модела на базата на разменни курсове, като изразява форуърдната инфлация чрез реални фактори и по такъв начин намалява броя на параметрите. Belgrade-Benhamou-Koehler дефинират индекса на потребителските цени посредством волатилността и корелацията.

ЛИТЕРАТУРА

1. Bårdsen, G., E. Jansen and R. Nymoen. Econometric Evaluation of the New Keynesian Phillips Curve. Oxford Bulletin of Economics and Statistics 2004, p. 27-46.
2. Bryan, M. and S. Cecchetti. The Seasonality of Consumer Prices. NBER Working Paper 1995.
3. Buys-Ballot, C. Les changements péridiques de températures. Utrecht 1847, Publication du Bureau central météorologique du France.

[49] Bryan, M. and S. Cecchetti. The Seasonality of Consumer Prices. NBER Working Paper 1995.

4. Calvo, G. Staggered Prices in a Utility Maximizing Framework. Journal of Monetary Economics 12/1983, p. 383-398.
5. Eitrheim, Ø. Testing the Role of Money in the Inflation Process. Norges Bank 2003.
6. Engle, R. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 50/1982, p. 984-1007.
7. Engle, R. and C. Granger. Cointegration ind Error-correction: Representation, Estimation and Testing. Econometrica 50/1987, p. 251-276.
8. Fisher, I. The Purchasing Power of Money: Its Determination and Relation to Credit Interest and Crisis. New York 2006.
9. Friedman, M. The Quantity Theory of Money: A Restatement. B: Studies in the Quantity Theory of Money. Chicago 1956.
10. Galì, J. and M. Gertler. Inflation Dynamics: A Structural Econometric Analysis. Journal of Monetary Economics 44/1999, p. 233-258.
11. Hallman, J., R. Porter and D. Small: Is the Price Level Tied to the M2 Monetary Aggregate in the Long Run? American Economic Review 81/1991, p. 841-858.
12. Ho, T. and S. Lee. Term Structure Movements and Pricing Interest Rate Contingent Claims. Journal of Finance 41/1986, p. 1011-1029.
13. Hull, J. and A. White. One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities. Journal of Financial and Quantitative Analysis 28/1993, p. 235-254.
14. Jacobson, T., J. Lyhagen, R. Larsson and M. Nessén. Inflation, Exchange Rates and PPP in a Multivariate Panel Cointegration Model. Working Paper Sveriges Riksbank 2002.
15. Johansen, S. and K. Juselius. Controlling Inflation in a Cointegrated Vector Autoregressive Model with an Application to U.S. Data. University of Copenhagen Working Paper 2003.
16. Juselius, K. Domestic and Foreign Effects on Prices and the Efficiency of Single-Equation Analysis. Journal of Econometrics 52/1992, p. 389-402.
17. Keynes, J. The General Theory of Employment, Interest and Money, Reprint 1997.
18. Koehler, E. and N. Belgrade. Evaluation de produits liés à l'inflation. Paris 2002; Beletski, T. Mathematical Models for Inflation and Pricing of Macro Derivatives. Kaiserslautern 2004.
19. Mercurio, F. Pricing Inflation-Indexed Derivatives. Working Paper, Milano 2004.
20. Nelson, D. ARCH Models and Diffusion Approximates. Journal of Econometrics 45/1990, p. 7-38.
21. Pigou, A. The Economics of Welfare. New York 2005.
22. Rudd, J. and K. Whelan. New Tests of the New Keynesian Phillips Curve. FRB Working Paper 2004.
23. Smets, F. and R. Wouters. An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area. Working Paper European Central Bank, Frankfurt/M. 2002.
24. Vasicek, O. An Equilibrium Characterisation of the Term Structure. Journal of Financial Economics 5/1977, p. 177-188.