

СЪДЪРЖАНИЕ

CONTENTS

Информационни технологии**I Information Technologies****1 Управление на разклонения в асинхронни микроконвейери***Димитър С. Тянев***1. Branch management in asynchronous micro-pipelines**3 *Dimitar S. Tyanov***2. Генериране на осветеност на контурни изображения посредством автовълнов процес в дискретна среда***Огнян И. Железов***2. Lighting Objects Using Autowave Propagation in Discrete Media**13 *Ognyan I. Zhelezov***3. Метод контроля для повышения достоверности рабочего диагностирования в вычислительных устройствах***Маоэл Said, Александр Дрозд, Мирослав Дрозд, Геннадий Дмитриев***3. Checking method to increase reliability of on-line testing in computing devices**18 *Maoel Said, Alexander Drozd, Miroslav Drozd, Gennady Dmitriev***4. Методи за съкращаване на излишните неврони от скрития слой на невронни мрежи с радиално-басисни функции***Ганка П. Ковачева***4. Pruning methods for hidden layer of Radial Basis Function Neural Networks**24 *Ganka P. Kovacheva***5. Обобщенная модель беспроводных каналов стандарта ieee 802.11***Сергей А. Нестеренко, Сергей .С. Нестеренко, Юлия С. Нестеренко***5. The Generalized Model of Wireless Channels Standard IEEE 802.11**30 *Sergej A. Nesterenko, Sergej S. Nesterenko, Ulja S. Nesterenko***6. Локализация квазипериодических последовательностей на изображениях***Светлана Г. Антощук, Анатолий А. Николенко, Тиен Т. К. Нгуен***6. Localization of Quasi Periodic Sequences at the Images**35 *Svetlana G. Antoshchuk, Anatoly A. Nikolenko, Tien T. K. Nguyen*

СЪДЪРЖАНИЕ

CONTENTS

7. Ключови генератори за потокови шифри на базата на преместващи регистри с линейни обратни връзки*Петър Ц. Антонов* 42 *Peter Ts. Antonov***Електронно обучение II E-Learning****8. Дизайн на Web базирана учебна среда способстващ освояването на знанията***Бойка Ж. Градинарова* 46 *Boyka Gradinarova***9. Спецификация IMS QTI и анализ тестовых заданий с коротким ответом на основе использования регулярных выражений***Валерий И. Давыдов, Петр Н. Фомин***10. Уровневая модель взаимодействий в электронном обучении***Вера В. Любченко* 59 *Vera V. Liubchenko***11. Изисквания за оформяне на статиите****7. Key Generators for Stream Ciphers on the Basis of Linear Feedback Shift Registers***Peter Ts. Antonov***8. Instructional Design of Web-Based Learning Environments***Boyka Gradinarova***9. Specification of IMS QTI and Analysis Test Tasks With a Short Answer on the Basis of the Use of Regular Expressions***Valery I. Davyдов, Peter N. Fomin***10. Layer Model of e-Learning Interactions***Vera V. Liubchenko***11. Guidelines for Preparing Manuscripts**

УПРАВЛЕНИЕ НА РАЗКЛОНЕНИЯ В АСИНХРОННИ МИКРОКОНВЕЙЕРИ

Димитър С. Тянев

Резюме: Разглеждат се структурни проблеми при синтеза на микроконвейери, предназначени да реализират алгоритми с условни преходи. Разглежданите структури изискват преопределение на понятието микроконвейер. В резултат на анализа на този общ проблем са изявени, анализирани и описани четири нови научни задачи, необходими за комплексното му решаване. Статията представя решението само на една от задачите, а именно синтезът на микроконвейерен автомат, управляващ звеното, което генерира стойността на условието за преход. Решението на тази задача се отнася до използване на конвейерни автомати с двуфазов протокол за трансфер на данни.

Ключови думи: Микроконвейери, Условни преходи, 2-фазов асинхронен автомат.

Branch management in asynchronous micro-pipelines

Dimitar S. Tyanov

Abstract: The paper considers structural problems in synthesis of micro-pipelines which implement algorithms with conditional jumps. These structures require pre-definition of the term “micro-pipeline”. As a result there are defined, analyzed and described four new scientific tasks necessary for solving this common problem. The paper presents the decision of only one of the tasks – synthesis of micro-pipeline that controls section generating value of the condition for jump. The solution of this problem refers to usage of pipeline automat with two-phase protocol for data transfer.

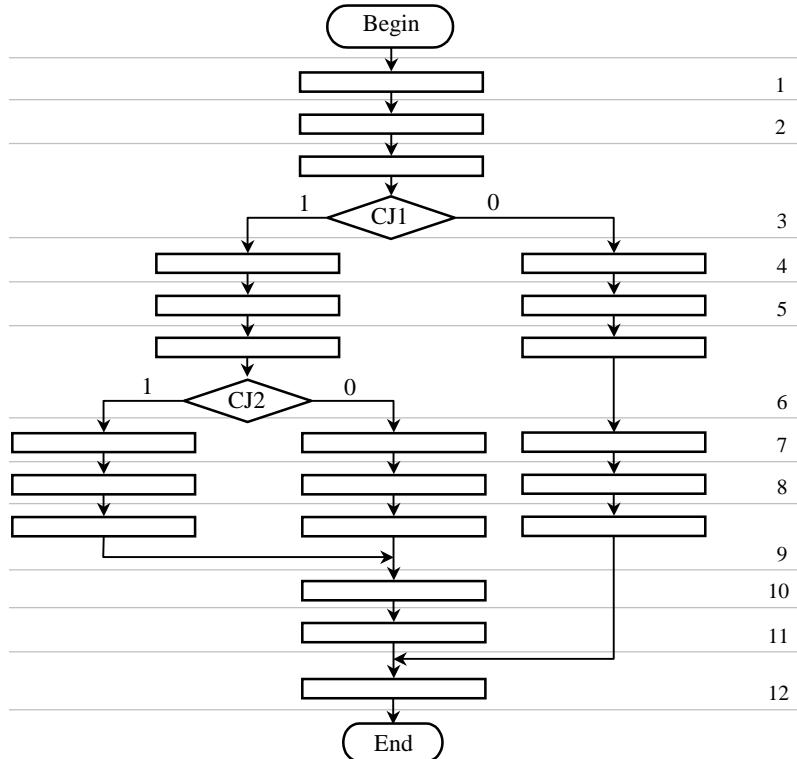
Keywords: Micro-pipelines, Branches, Two-phase asynchronous automat.

I. Въведение

Разглеждаме представената на фигура 1 структура на примерен алгоритъм, който е аппаратно реализиран и чието изпълнение е конвейерно организирано. Има се предвид още, че този алгоритъм е детализиран и неговите изпълними блокове са реализирани аппаратно с помощта на методите, изложени в [10], [11], [12], [13], [14] и [18]. Всеки изпълним блок от блок-схемата представлява отделно еднотактово или многотактово микроконвейерно звено, според определенията изложени в [16] и [19]. Това означава, че при интерпретацията на алгоритъма, от гледна точка на изпълнимите операции, всеки изпълним блок може да бъде разглеждан най-общо като многотактово звено. Всяко многотактово звено обаче от своя страна може да представлява в своята вътрешност и по-сложна съставна структура от произволни микроконвейерни звена, имаща вид, подобен на показаната.

Както се вижда, представеният алгоритъм съдържа няколко линейни участъка, но като цяло може да се определи като разклонен, въпреки изказаните по-горе съображения. Двете условия за преход CJ (*conditional jump*) формират следните възможни пътища за изчислителния процес:

1. Begin; 1; 2; 3 (CJ1=true); 4; 5; 6 (CJ2=true); 7; 8; 9; 10; 11; 12; End .
2. Begin; 1; 2; 3 (CJ1=true); 4; 5; 6 (CJ2=false); 7; 8; 9; 10; 11; 12; End .
3. Begin; 1; 2; 3 (CJ1=false); 4; 5; 6; 7; 8; 9; 12; End .



Фиг.1 Структура на примерен алгоритъм

Като се има предвид, че всеки алгоритмичен път *Begin-End* е уникален, съответните последователни микроконвейерни звена в паралелните клонове са поставени в едно и също поредно ниво на конвейера, за който тук те са общо 12. В нива от 4 до 9, в които има по няколко микроконвейерни звена, в момента на всяко отделно изпълнение, работи само съответно включеното в дадения алгоритмичен път звено.

Общо разбиране за понятието микроконвейер

Така поставеният проблем за хардуерна реализация на алгоритъм с условни преходи изисква ново разбиране за понятието микроконвейер, което следва да поясним.

В условията на конвейерна организация данновите връзки между конвейерните регистри се управляват от конвейерни автомати. Това означава, че превключването на всеки конвейерен автомат в ниво с разклонения е в зависимост от стойността на съответното условие за преход. Като продължение на микроконвейера, след нивото, генериращо условие за преход, се налага микроконвейерна реализация и на двата алгоритмични клона. Така понятието микроконвейер формално (привидно) се нарушава, тъй като се нарушава общо приетото разбиране за последователно подреждане на микроконвейерни звена. В разглеждания случай обаче присъствието на физическата (апаратната) реализация на всички възможни изчислителни пътеки *Begin-End* е неизбежно и е следствие от свойството общност на всеки алгоритъм [29]. Колкото и много да са те обаче, всеки един път, по който се развива изчислителният процес на дадена задача, е уникален по това, че преминава последователно през последователно активирани микроконвейерни звена и точно в този смисъл, точно тези звена образуват една последователна верига, която напълно съответства на понятието микроконвейер. Ето защо, присъствието в схемата на други аппаратно реализирани паралелни пътища, не противоречи на общо приетото разбиране за микроконвейер, и като цяло всяка подобна структура, с по-висока степен на общност, също може да се определи като микроконвейер. С други думи, може да се приема, че структурата е съставена от няколко конвейера, чито общи части са обединени, така че тя има едно начало и един край.

II. Нови задачи

Микроконвейерната реализация на подобен род общи алгоритмични структури среща нов и многостранен проблем – управление хода на изчислителния процес при алтернативни условия. Анализът на този проблем изявява няколко нови и нерешавани задачи. Анализът и изявяването на тези нови задачи, свързани с приетата обща постановка, ние приемаме като самостоятелен резултат от нашите изследвания.

На първо място очевидна е задачата за синтез на конвейерен автомат в точката на условния преход. Този автомат съществено се различава от обикновения линеен автомат по това, че той трябва да реши задачата за избор на един от двата алгоритмични клона. Тази задача е неизбежно свързана със следващите. Всички нови задачи са свързани с проектиране на конвейерни автомати в особените точки на алгоритмичните структури от общ вид.

В трудовете на някои автори [1], [2], [3], [4], [5], [6], [7], [8], [9] се коментира възможността за управление на разклонения в микроконвейерите, но единствено поради паралелно дублиране на линейни конвейерни участъци, които при това са напълно еднакви по функционалност. Разклонения поради условни преходи в посочените микроконвейерни реализации не се съдържат, а така също липсва систематично и общо изследване на този проблем.

Оригинални конвейерни автомати, управляващи трансфера между еднотактови, многотактови и смесени по тип микроконвейерни звена ние сме предложили в [16], [17] и [19]. Общото между тях е това, че тези автомати, в цялата им разновидност, поддържат само микроконвейери с линейна структура, каквито са и множеството структури, разглеждани в публичното пространство, начиная с основополагация труд [20]. Апаратната реализация на микроконвейерно звено с условен алгоритмичен преход ние вече сме разглеждали в [10], където е изложен анализът на проблема и са предложени два варианта за неговата реализация. Общото между тях е това, че стойността на логическото условие за преход CJ се използва в звеното за управление на данновата шина, за да бъдат резултатите насочени към актуалното разклонение, реализирани по този начин конкретен алгоритмичен път *Begin-End*. Казаното току що съответства на нива 4 и 7 в схемата от фигура 1 (обърнете внимание на мястото на линиите, които ограничават състава на звената). Формулираната в [10] задача за управление на трансфера на данни в микроконвейер с разклонения е също нова, а нейното общо решение ще бъде изложено тук. Същността на така поставената задача, се състои в синтез на конвейерен автомат, управляващ микроконвейерното звено, генериращо условието за преход. Логическата стойност на това условие трябва да реши в кой от клоновете трябва да постъпи заявката *Request*. С тази задача е свързана задачата за приемане на сигнала потвърждение *Acknowledgement* от съответното звено, приело заявката. Поради неразривната същност на сигналите *Req* и *Ack* ние приемаме задачата за една, въпреки двете отделни решения, които тя притежава.

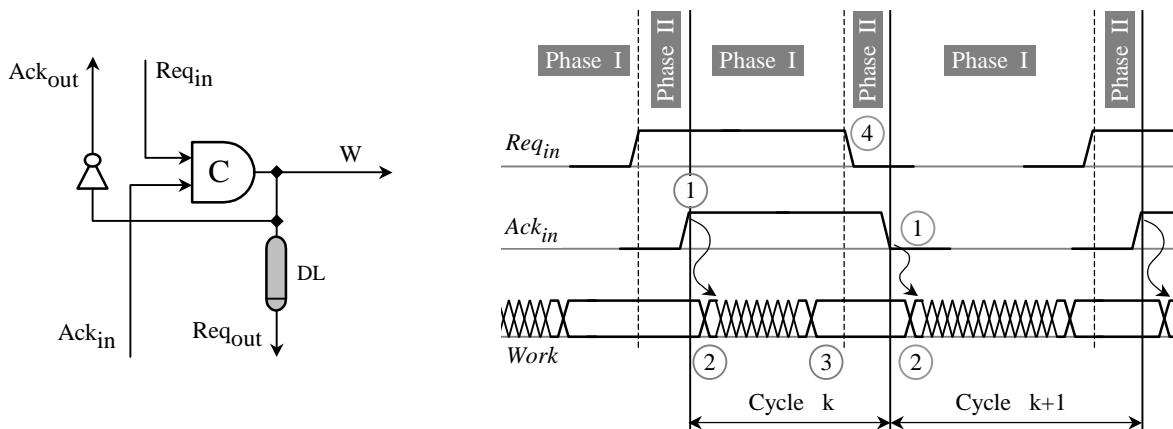
На второ място е задачата за синтез на конвейерен автомат, който може да управлява микроконвейерни звена в общите точки на алгоритъма. Както може да се види от фигура 1, в ниво 9 и в ниво 12, входните точки на съответните микроконвейерни звена обединяват повече от един изход от предходни звена. Постъпването на входа на дадено звено на готови резултати, идващи от няколко предходни звена, разположени топологически паралелно, се изявява като нов и самостоятелен проблем. Същността на неговото решение, се състои в синтез на конвейерен автомат, управляващ приемащото микроконвейерно звено при паралелно постъпване във времето в общия случай на повече от една заявка *Request*. В огледален смисъл с тази задача е свързана още една задача – задачата за генериране на сигнал потвърждение *Acknowledgement* към съответно свързаното предходно звено. Поради неразривната същност на сигналите *Req* и *Ack* и тук приемаме задачата за една, но тя също ще притежава две отделни решения.

Постъпващите към приемащото звено в общата точка заявки от паралелни предходни звена повдига третата задача, същността на която е изборът на заявка. За решаване на задачата за избор на заявка и приемане на съответните данни, конвейерният автомат на приемащото звено трябва да изпълни арбитрираща процедура. Реализацията на тази процедура представлява третата по ред задача. Задачата за арбитриране на заявки по принцип е известна и за нея са познати различни аппаратни решения [29], но в условията на микроконвейерно управление я приемаме за нова.

И на края четвъртата задача: в общата точка, където се обединяват няколко разклонения, заявките, които придвижват получените в тези клонове данни, се отнасят до различни задачи, стартирани преди това в микроконвейера. Редът, в който съответните им междинни резултати достигат точката на обединяване, съвсем не може да се очаква да съответства напълно на реда, в който са стартирани задачите, на които те принадлежат. С други думи, в приемащото звено, което стои в общата точка, едва ли ще постъпват данни в правилния ред. Така наличието на разклонения в микроконвейера води до проблем, същността на който се състои в това, че излизашите на изхода на конвейера окончателни резултати едва ли ще бъдат в онзи ред, който съответства на реда на стартиране. Тук явно трябва да се формулира нова, четвърта по ред задача, която изисква въвеждане на цялостна система за идентифициране на реда и съответствието на изходните резултати.

III. Микроконвейер с дву-фазов протокол за трансфер

Управлението на хода на изчислителния процес, както беше отбелязано, е задача на конвейерните автомати,eto защо същността на първата задача е проектиране на конвейерен автомат, зависим от условието за преход. Като първи вариант за решение на тази задача тук е разгледано управление с автомат, използващ асинхронен Мюлер C-елемент [20], [21]. Автоматите от този вид (фигура 2) реализират 2-фазов протокол за трансфер. Събитията в този протокол: заявката Req_{in} , потвърждението Ack_{in} и функционирането на микроконвейерното звено $Work$, за два последователни работни цикъла (k) и ($k+1$), са номерирани във времедиаграмата на протокола както следва:



Фиг. 2 Дву-фазов конвейерен автомат и протокол

1. Сигнал Ack_{in} се превключва в единица. При $Req_{in}=Ack_{in}=1$, конвейерният автомат на текущото звено се превключва в единица (състояние 1), генерирайки сигнал за запис на данни в неговия регистър фиксатор ;
2. В операционната логика след регистъра, в който току що са записани нови данни, стартира преходен процес – начало на изчисления в текущото звено ;
3. Изчисленията завършват – резултатът е готов ;

4. В това състояние конвейерният автомат на текущото звено очаква от предходното звено превключването на съгнала Req_{in} в ниско ниво ($Req_{in}=0$), както и потвърждението за приети данни от следващото звено ($Ack_{in}=0$) ;

Следващият ($k+1$) трансферен цикъл започва със същите действия:

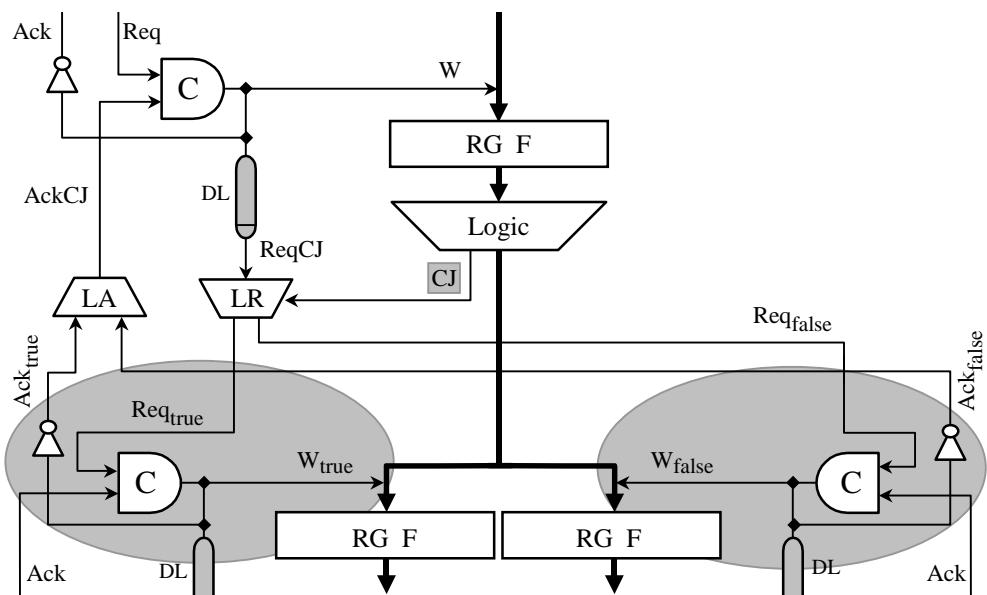
1. В момента, в който се постигне равенството ($Req_{in}=Ack_{in}=0$), започва нов цикъл.

Конвейерният автомат на текущото звено се превключва обратно в ниско ниво (състояние 0), генерирачи сигнал за запис на данни в неговия регистър ;

2. Записът на данни е извършен и в текущото звено стартира нов цикъл на изчисления.

Времедиаграмата позволява да бъде съобразено, че четирите възможни комбинации от стойностите на входните сигнали (Req_{in}, Ack_{in}) са поделени по равно между двете състояния на конвейерния автомат – комбинации 11, 01 за състояние 1, и комбинации 00, 10 за състояние 0.

На фигура 3 е представена част от структура на микроконвейер, съдържаща звено, което генерира условие за преход CJ. Към логическата схема на конвейерния автомат, управляващ това звено, се отнасят две допълнителни логически схеми, чийто синтез ще представим по-долу: LA – схема за формиране на сигнала $AckCJ$, и LR – схема за генериране на актуалните заявки.



Фиг. 3 Структура с конвейерно разклонение

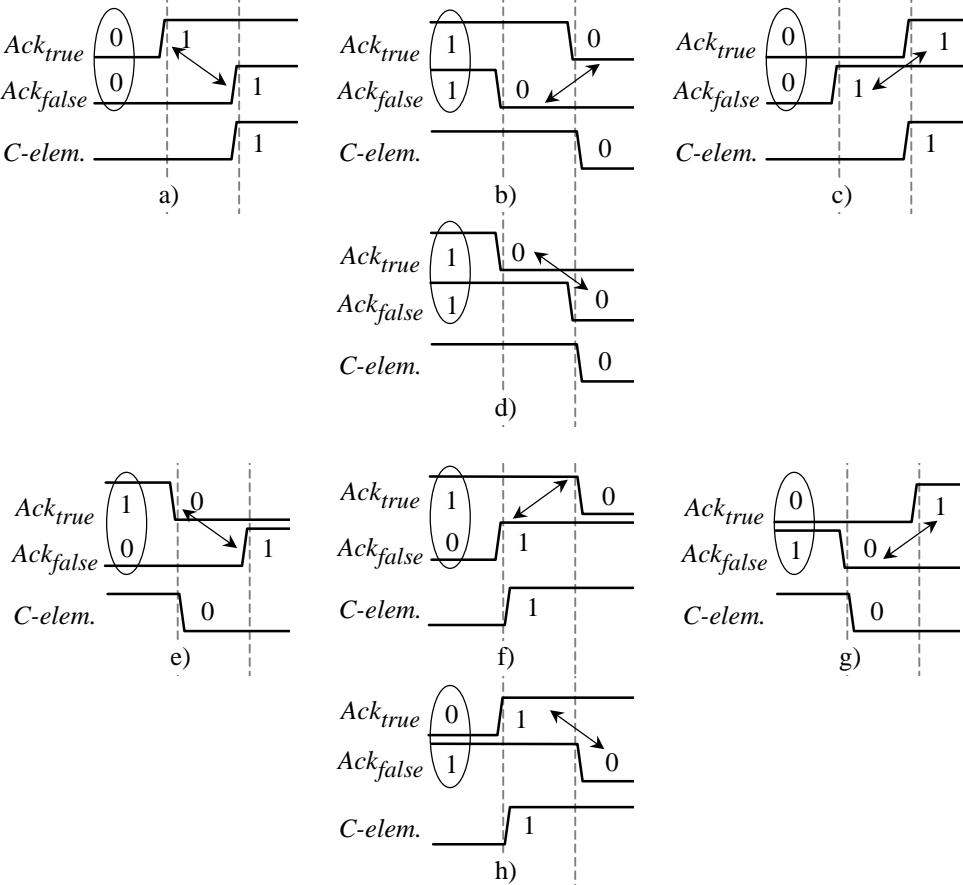
В структурата са показани и двата алгоритмични клона. Който и от клоновете да бъде избран при конкретното изпълнение на изчисленията, следва да сработи онзи начален конвейерен автомат, който е определен от актуалната стойност на логическото условие за преход CJ. В същност, в началото на всяко от разклоненията стои регистър фиксатор RG_F, но микрооперация запис по сигнал W ще изпълни само един от тези регистри.

Формиране на потвърждение към автомата с условен преход

Превключването на конвейерният автомат в звеното с условен преход е функция от два сигнала: Ack_{true} и Ack_{false} , оповестяващи готовността на всеки от клоновете по отделно, т.е. те са "родители" на сигнала $AckCJ$. Спрямо звеното с условен преход, сигналите Ack_{true} и Ack_{false} са еднакво паритетни, но във времето те са конкурентни. Последното означава, че във времето моментите на тяхно поредно превключване могат и да съвпаднат, при което самите превключвания могат да бъдат както еднакви, така и противоположни. Трябва да се отбележи, че състезателността между тези превключвания е несъстоятелна като се има

предвид условието за переход CJ, чиято стойност ще бъде формирана в последствие, т.е. със закъснение. В крайна сметка процесът ще продължи в избрания клон, а той може да не е този, чийто сигнал Ack е спечелил състезанието.

Синтезът на логическата схема LA, която ще обедини сигналите Ack_{true} и Ack_{false} и ще формира функцията $AckCJ$, трябва да отчита всички възможни ситуации както между стойностите, така и във времето. За целта тези ситуации са изобразени на фигура 4.



Фиг. 4 Възможни ситуации при переход в сигналите Ack_{true} и Ack_{false}

Например, рисунка 4а) представя завареното състояние (00) на сигналите Ack_{true} и Ack_{false} , при което и двата се превключват в 1, но този от клон “истина” – Ack_{true} , постига това по-рано във времето. Рисунка 4с) представя същото изходно състояние (00), но във времето състезанието печели сигналът Ack_{false} .

Стойността на потвърждението $AckCJ$ се формира въз основа на разбирането, че микроконвейерното звено с условен переход може да бъде стартирано от своя автомат, ако поне един от клоновете е готов и е превключил стойността на своя сигнал за потвърждение. Този факт подлежи на запомняне, за да не се повлияе стойността на сигнала $AckCJ$ от превключването на закъснялото потвърждение, идващо от другия клон. Така изложеното разбиране обосновава Мюлер C-елемента като функция, която може да обедини сигналите Ack_{True} и Ack_{false} до сигнала $AckCJ$. По тази причина в рисунките от фигура 4 всички времедиаграми съдържат още изходната стойност на C-елемента като функция от двете входни потвърждения. В смисъла на разгледания вече 2-фазов протокол, всяко превключване на C-елемента (с други думи на сигнала $AckCJ$) ще стартира звеното с условен переход.

Генериране на заявки към разклоненията

Конвейерният автомат, управляващ звеното с условен переход, разпространява към следващите звена заявка, отбелязана като $ReqCJ$. За съжаление заявката не може да се подаде

директно към входовете на конвейерните автомати в началото на всеки от алтернативните клонове. Съответните запитвания, които следва да получат конвейерните автомати, са означени с Req_{true} за клона “истина” и с Req_{false} за клона “лъжа” и са функция на логическата схема LR (фигура 3). Непосредственото подключване на заявката $ReqCJ$ не е възможно, защото в точката на разклонението тя влиза в сложна функционална връзка с логическата стойност на условието CJ от една страна, и с текущото (завареното) състояние на конвейерния автомат в началото на всеки клон, от друга страна. Това пък се налага от вида на самите автомати, които, както вече беше споменато, използват 2-фазов протокол за управление на трансфера. Последното означава, че всяко тяхно превключване ($0 \rightarrow 1$ и $1 \rightarrow 0$) причинява запис в регистрите фиксатори и старт на изчисленията в звената. Ако се разглежда превключване в началото на клон от микроконвейера, то освен логическата стойност на условието за переход CJ (0 или 1) следва да се отчете и завареното състояние на съответния автомат. Следва да се обърне внимание на това, че C-елементът се превключва както при съчетаване на две входни единици, така и на две нули. С други думи сигналите Req_{true} и Req_{false} са функции не само от превключването на сигнала $ReqCJ$ и от стойността на условието CJ, но и от състоянието на конвейерните автомати на входа на разклоненията. Например, ако стойността на условието за переход е единица (“истина”), това означава, че изчисленията трябва да продължат в клона “истина” (левият клон на структурата от фигура 3). Ако състоянието на автомата в този клон е единица, което се поддържа във времето от $Req_{true}=1$, то той трябва да се превключи в състояние нула, за да стартира чрез заден фронт на сигнала W_{thru} тези изчисления. За да стане това, на входа на този C-елемент трябва да се съчетаят две нули. Не трябва да се забравя и фактът, че всяка нова стойност на сигнала $ReqCJ$ (и 0 и 1) представлява по същество нова заявка, излизаша от автомата в звеното на разклонение.

Логиката, която току що беше пояснена, е изразена чрез следните таблици на истинност, в които сигналите W_{true} и W_{false} представляват състоянието на съответните C-елементи.

Таблица 1 Заявка към конвейерен автомат в клон “истина”

CJ	W_{true}	ReqCJ	Req_{true}
0	0	появява се заден фронт $\downarrow 0$	0, не се превключва
0	0	появява се преден фронт $\uparrow 1$	0, не се превключва
0	1	появява се заден фронт $\downarrow 0$	1, не се превключва
0	1	появява се преден фронт $\uparrow 1$	1, не се превключва
1	0	появява се заден фронт $\downarrow 0$	$\uparrow 1$, превключва се
1	0	появява се преден фронт $\uparrow 1$	$\uparrow 1$, превключва се
1	1	появява се заден фронт $\downarrow 0$	$\downarrow 0$, превключва се
1	1	появява се преден фронт $\uparrow 1$	$\downarrow 0$, превключва се

Таблица 2 Заявка към конвейерен автомат в клон “лъжа”

CJ	W_{false}	ReqCJ	Req_{false}
0	0	появява се заден фронт $\downarrow 0$	$\uparrow 1$, превключва се
0	0	появява се преден фронт $\uparrow 1$	$\uparrow 1$, превключва се
0	1	появява се заден фронт $\downarrow 0$	$\downarrow 0$, превключва се
0	1	появява се преден фронт $\uparrow 1$	$\downarrow 0$, превключва се
1	0	появява се заден фронт $\downarrow 0$	0, не се превключва
1	0	появява се преден фронт $\uparrow 1$	0, не се превключва
1	1	появява се заден фронт $\downarrow 0$	1, не се превключва
1	1	появява се преден фронт $\uparrow 1$	1, не се превключва

Въз основа на горните таблици на истинност са синтезирани следните логически функции:

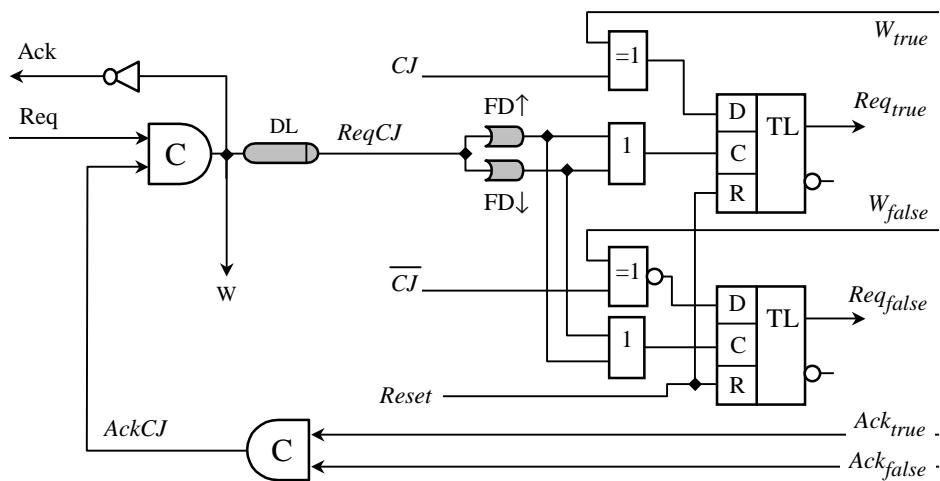
$$Req_{true} = CJ \oplus W_{true} . \quad (1)$$

$$Req_{false} = \overline{CJ \oplus W_{false}} . \quad (2)$$

Както се вижда, логиката на заявките Req_{true} и Req_{false} не зависи от заявката $ReqCJ$, което се очакваше. Зависимостта на заявките Req_{true} и Req_{false} не е от стойността на сигнала $ReqCJ$, а от времето, т.е. от моментите на превключване на сигнала $ReqCJ$. Следва да се разбира, че смяната на стойността на $ReqCJ$, т.е. появата на негов фронт, маркира момента във времето, в който операционната логика на това звено завършва изчисленията си. Този момент не е задължително да съвпада с появата на истинната стойност на условието CJ. В зависимост от сложността на изчислението на CJ, в общия случай следва да се приема, че истинната стойност на CJ може да се появи и по-рано във времето спрямо новия фронт на сигнала $ReqCJ$ или най-късно едновременно с него и никога по-късно от него. В една непосредствената реализация на (1) и (2), по-ранното явяване на CJ ще доведе до по-ранно формиране на стойностите на заявките Req_{true} и Req_{false} , а от там и до по-ранно стартиране на съответния клон на конвейера. Това стартиране ще започне със запис на данни в регистъра фиксатор, а те в общи случаи все още няма да са достигнали във времето истинните стойности. Така най-вероятно изчисленията могат да стартират с грешни данни.

Основният извод, който се налага от изложените съображения, е, че формулите (1) и (2) определят стойностите на заявките, но моментът, в който те следва да се появят и да подействат, се определя от момента на превключване на сигнала $ReqCJ$. Или казано по друг начин, новите стойности на заявките Req_{true} и Req_{false} трябва да се явят в отговор на фронт в сигнала $ReqCJ$. Последното означава, че формирането на заявките Req_{true} и Req_{false} , не е възможно да бъде постигнато само с чиста комбинационна логика.

Изложените разсъждения доказват, че времевата зависимост на изчислениите стойности (1) и (2) може да се реализира само с помощта на запомнящ елемент – тригер. Тъй като запис трябва да се извърши при всяко превключване на C-елемента, синхронизиращият тригер трябва да бъде от тип DEDTFF – [23], [24], [25], [26], [27], [28] и др. (D-тригер, работещ и по двата фронта). В окончателната логическа схема на конвейерния автомат (фигура 5) е представено едно предпочитано от нас решение, основано на обикновен D-Latch тригер и два детектора на фронт [22], [23], [24] – $FD\uparrow$ за преден и $FD\downarrow$ за заден. Тя съдържа C-елемента, обединяващ потвържденията Ack_{true} и Ack_{false} , както и двете тригерни схеми, генериращи заявките Req_{true} и Req_{false} .



Фиг. 5 Схема за генериране на заявка в клон “истина”

Както се вижда от схемата, пулс-генераторите се обединяват в схема ИЛИ и реализират запис по входа С на тригера. Изчислената според (1) стойност на заявката Req_{true} постъпва по входа D и се съхранява в тригера до следващия път, когато ще бъде избран същия клон. Сигналът *Reset* е необходим в началния момент, в който всички конвейерни автомати принудително се установяват в изходно състояние. Аналогична логическа схема формира според (2) заявката Req_{false} към конвейерния автомат в клон “льжа”.

IV. Заключение

Въпреки че разгледаната тук задача получи своето решение, в цялост проблемът с микроконвейерната реализация на общи алгоритмични структури не беше решен. Той може да се счита за решен при наличие на решения и за останалите задачи, формулирани в началото. Тези четири задачи са неразривно свързани и нямат практическа самостоятелност. Възможността за проектиране на изчислителен процес, който съдържа разнообразни алгоритмични структури чрез методите на микроконвейерна организация позволява значително повишение на производителността му. Това е резултат от една страна благодарение на възможността за хардуерна реализация, а от друга страна благодарение на самата конвейерна организация, която е основен метод за внедряване на паралелизъм в изчисленията. Като имаме предвид възможностите на съвременната елементна база, в лицето на програмируемите ГИС от типа на FPGA, определено считаме, че са налице необходимите условия за реализация на предложения тук метод.

Литература

- [1]. Feng Shi, Makris; Yiorgos, Nowick, Steven M.; Singh, Montek; *Test Generation for Ultra-High-Speed Asynchronous Pipelines*, International Test Conference, p. 39.1–39.10, Nov 2005.
- [2]. Singh, Montek; Nowick, Steven M.; *MOUSETRAP: Ultra-High-Speed Transition-Signaling Asynchronous Pipelines*, Proc. In. Conf. Computer Design (ICCD), Nov. 2001, pp. 9-17.
- [3]. Singh, Montek; Nowick, Steven M.; *MOUSETRAP: Designing High-Speed Asynchronous Digital Pipelines*, <http://www.cs.columbia.edu/~nowick/columbia-cisl-seminar-mousetrap-pt2.pdf>.
- [4]. Ozdag, Recep O.; Singh, Montek; Beerel, Peter A.; Nowick, Steven M.; *High-Speed Non-Linear Asynchronous Pipelines*, http://www.date-conference.com/archive/conference/proceedings/papers/2002/date02/pdffiles/10b_1.pdf.
- [5] Davis, A.; Nowick, S. M.; *An Introduction to Asynchronous Circuit Design*, Technical Report UUCS-97-013, Computer Science Department, University of Utah, Sep., 1997. <http://www.cs.columbia.edu/~nowick/ald-nowick-tr-intro.pdf> .
- [6]. Frank te Beest, *Click Elements - An implementation style for data-driven compilation*, http://journals2.scholarsportal.info/details.xqy?uri=/15228681/v2010inone/3_ceaisfdc.xml .
- [7]. Fawaz, Khodor, Arslan, Tughrul, Lindsay, Iain, *Conditional Acknowledge Synchronisation in Asynchronous Interconnect Switch Design*, 2009 NASA/ESA Conference on Adaptive Hardware and Systems, Issue Date:July 2009, pp. 126-131.
- [8]. Fu-Chiung Cheng; Shu-Ming Chang; Chi-Huam Shieh; *Detection and Generation of Self-Timed Pipelines from High Level Specifications*, 20th International Conference on VLSI Design (VLSID'07), January 2007, pp. 413-418.
- [9]. Beerel, Peter A.; Ozdag, Recep O.; Ferretti, Marcos; *A Designer's Guide to Asynchronous VLSI*, ISBN 978-0-521-87244-7, Cambridge University Press, 2010.
- [10]. Tyanov, Dimitar; Josifov, Veselin; Kolev, Stamen; *Operational structures without controlling automata*, International Workshop on Network and GRID Infrastructures, 27-28 Sept 2007, Bulgarian Academy of Sciences, Sofia, Bulgaria.

- [11]. Тянев, Д.С.; Колев, С.И.; *Метод за реализация на апаратни самоуправляващи се циклически структури*, Годишник на ТУ-Варна, Юбилеен сборник “45 години ТУ-Варна”, 2007, ISSN 1311-896X, стр. 130-135.
- [12]. Тянев, Д.С.; Колев, С.И.; Янев, Д.В.; *Метод за реализация на апаратни самоуправляващи се циклически структури - част II*, Списание “Компютърни науки и технологии”, ТУ-Варна, ISSN 1312-3335, година V, брой №2/2007, стр. 23-30.
- [13]. Tyanev, Dimitar; Kolev, Stamen; Yanev, Dragomir; *Micro-pipeline Section For Condition-Controlled Loop*, International Conference on Computer Systems and Technologies - *CompSysTech'09*, 18-19 June 2009, Ruse, Bulgaria, pp. I.4 1-5.
- [14]. Dimitar Tyanev, Dragomir Yanev, Stamen Kolev, *Method for realization of self-controlling loop apparatus structures*, Fifth International Scientific Conference *Computer Science'2009*, 5-6 November 2009, Sofia, Bulgaria.
- [15]. Tyanev, Dimitar; *Principle scheme of aperiodic finite state machine*, Списание “Комп. науки и технологии”, ТУ-Варна, ISSN 1312-3335, г. V, брой №2/2008, стр. 3-8.
- [16]. Tyanev, Dimitar; *Four-phase micro-pipeline with one-cycle and multi-cycle micro-pipeline sections*, Списание “Компютърни науки и технологии”, ТУ-Варна, ISSN 1312-3335, година VII, брой №1/2009, стр. 3-12.
- [17]. Tyanev, Dimitar; Popova, Stefka; *Asynchronous micro-pipeline with multi-stage sections*, *ICEST'2010*, 23-26 June 2010, Ohrid, Macedonia.
- [18]. Tyanev, Dimitar; Kolev, Stamen; Yanev, Dragomir; *Race Condition free Asynchronous Micro-Pipeline Units*, International Conference on Computer Systems and Technologies - *CompSysTech'10*, 17-18 June 2010, Sofia, Bulgaria.
- [19]. Kolev, Stamen; Tyanev, Dimitar; *Early set to zero micropipeline*, International Conference on Computer Systems and Technologies - *CompSysTech'10*, 17-18 June 2010, Sofia, Bulgaria.
- [20]. Sutherland, Ivan E.; *Micropipelines*, <http://www.jdl.ac.cn/turing/pdf/p720-sutherland.pdf>.
- [21]. Миллер, Реймонд Е.; *Теория переключательных схем, том 2 – (последовательностные схемы и машины)*, Москва, Издательство “Наука”, 1971.
- [22]. Holdsworth, Braian; Woods, Clive; *Digital Logic Design*, Fourth Edition, Newnes Imprint of Elsevier, ISBN: 8131200620, 2005.
- [23]. Breeding, Kenneth J.; *Digital Design Fundamentals*, Second Edition, Prentice Hall, ISBN-10: 0132112779, 1997.
- [24]. Kumar, Rajesh; Bollapalli, Kalyana C.; et al.; *A Robust Pulsed Flip-flop and its use in Enhanced Scan Design*, <http://iccd.et.tudelft.nl/2009/proceedings/97Kumar.pdf> .
- [25]. Tschanz, James W.; *Flip-flop circuit having dual-edge triggered pulse generator*, United States Patent №US6608513, Aug., 19, 2003.
- [26]. Mohideen, S. Kaja; J. Perinbam, Rajapaul; *Design of Low Power Double Edge Triggered D Flip Flop*, Asian Jornal of Information Techology 5(10): 1113-1116, 2006, <http://www.medwelljournals.com/abstract/?doi=ajit.2006.1113.1116> .
- [27]. Pedram, Massaud; Qing Wu, Xunwei Wu, *A New Design for Double Edge Triggered Flip-flops*, Proc. ASP-DAC, Yokohama, 1998, <http://atrak.usc.edu/~massoud/Papers/det-ff.pdf> .
- [28]. Takashi Asano, Kouichi Yamada, *Double Edge Triggered Flip-flop circuit*, US Patent №20100034030, 11 Feb.2010, <http://www.freshpatents.com/dt20100211ptan20100034030.php>.
- [29]. Тянев, Димитър; *Организация на компютър*, том 2, Технически Университет - Варна, ISBN 978-954-20-0413-4, 2008.

За контакти:

Доц. д-р инж. Димитър Тянев
 Катедра “Компютърни науки и технологии”
 Технически университет - Варна
 e-mail: dstyanev@yahoo.com
 Рецензент: доц. д-р инж. Н. Рускова, ТУ – Варна

ГЕНЕРИРАНЕ НА ОСВЕТЕНОСТ НА КОНТУРНИ ИЗОБРАЖЕНИЯ ПОСРЕДСТВОМ АВТОВЪЛНОВ ПРОЦЕС В ДИСКРЕТНА СРЕДА

Огнян И. Железов

Резюме: В тази статия се предлага алгоритъм за генериране на осветеност на контурни изображения посредством автовълнов процес (АВП) в дискретна среда. Използва се модификация на бинарния АВП, като растерното изображение се представя като мрежа от дискретни елементи, всеки от които може да бъде в едно от N възможни състояния. Предлаганият модел на АВП реализира еднопосочко разпространение на вълновия процес на променяне на състоянието на елементите. На следващата стъпка от алгоритъма от състоянието на елементите се получава яркостта на точките на изходното изображение. Като резултат се получава изменение на яркостта на точките от вътрешността на контурните обекти, пропорционално на разстоянието от контура на обекта. Алгоритъмът притежава предимствата, характерни за алгоритмите, използвани АВП – високо бързодействие и възможност за паралелно изпълнение на операциите.

Lighting Objects Using Autowave Propagation in Discrete Media

Ognyan I. Zhelezov

Abstract: In this paper is proposed one algorithm for lighting of contour images using autowave propagation (AWP) in excitable media. Algorithm uses one modification of discrete AWP. Source rastered image is described as a net of discrete elements. Propagation of AWP within this media is a process of switching of elements from state 255 (white) to state, obtained from the brightness of neighbor elements, having color, different from white. As a result, the value of pixel's brightness inside an objects contour is changed as a function of distance from pixel (AWP element) to contour curve. Algorithm has the advantages of AWP algorithms – fast performing and possibility parallel execution.

Key words: autowaves propagation, lighting of objects

1. Увод

Генерирането на осветеност на двумерни изображения с цел да се получи илюзия за тримерност, по правило се извършва в интерактивен режим. Графичните 3D редактори предоставят функции, чрез които може да се преобразуват двумерни изображения в определени видове тримерни и като резултат да се получи тяхното фотореалистично осветяване. В статията се предлага алгоритъм за генериране на осветяване на двумерни контурни изображения посредством автовълнов процес (АВП), който се изпълнява с по-малък брой операции при определени ограничения на функцията на изменение на яркостта и на положението на източника на осветяване.

2. Общи сведения за автовълновите процеси

Автовълновите процеси (АВП) са процеси на разпространение на нелинейни вълни във възбудима среда. Те се описват посредством нелинейни диференциални уравнения от вида [1]

$$\frac{\partial u}{\partial t} = D \cdot \frac{\partial^2 u}{\partial x^2} + f(u) \quad (1)$$

Съществуват различни модели на разпространение на АВП в дискретна среда. При най-прости бинарен модел всеки елемент от дискретната среда има само две възможни състояния (могат да бъдат обозначени условно с нула и единица). Разпространението на вълната при бинарен АВП се извършва по следния алгоритъм:

Ако за даден такт от дискретното време един елемент е в състояние 0 и има поне един съседен елемент, който е в състояние 1, то по време на следващия такт той превключва в състояние 1. В противен случай елементът запазва състоянието си. Дори този най-прост модел на АВП дава възможност за извършване на ефективна паралелна обработка на бинарни изображения – възстановяване на прекъснати контури, изглаждане на контури, скелетизация и др.

В доклада се предлага алгоритъм за генериране на осветеност на контурни изображения посредством АВП в дискретна среда с 256 възможни състояния на елементите, съответстващи на 8-битово кодиране на яркостта на точките от изображението. Подобренето на алгоритъма, предложен в [4] е във въвеждането на функцията на осветяване $F_b(n)$, която задава яркостта на пиксел от вътрешността на обекта като функция на състоянието на съответния елемент от дискретната среда, а не като функция на разстоянието от контура на обекта до пиксела.

3. Формулировка на задачата за генериране на осветеност

Генерирането на двумерни изображения със зададена форма по правило започва с получаването на контурно изображение – изображение, което съдържа само контурни криви, които отделят обекта от фона на изображението. Генерирането на осветеност на обект, представен с контурна крива създава излюзия за тримерност, което подобрява вида на изображението.

Ако се приеме, че контурната крива на обекта е получена от двумерна проекция на тримерен обект, то очевидно тя не съдържа достатъчно информация, по която да може да бъде възстановен тримерният обект. Известната в теорията на цифровата обработка на изображения теорема за *проекционния срез* доказва, че едно тримерно изображение може да бъде възстановено напълно чрез определен брой проекции, получени за проекционни ъгли, равномерно разположени в интервала от 0 до 2π . В случай, че се разполага само с контурно изображение, получено от една двумерна проекция, може чрез него да се получи тримерен обект (и съответно осветеност на обекта) само ако се направят определени предположения за формата на обекта. Съвременните програмни средства за тримерна графика (3DMAX, OpenGL, DirectX) предлагат две основни функции за преобразуване на контурно изображение в тримерен обект – изтегляне (екструдер) и ротация. Екструдерът преобразува контурната крива в криволинеен цилиндър, ограничен от контурната крива и сечението на две успоредни равнини. Ротацията на контурната крива около избрана ос създава тримерен обект с кръгова симетрия. Полученият тримерен обект се описва чрез координатите на върховете на мрежа от полигони, като осветяването на всеки полигон се изчислява чрез осветеността на неговите върхове. Тези алгоритми за генериране на тримерни обекти и осветяването им, въпреки че гарантират получаването на фотореалистични обекти със зададени геометрични характеристики и осветеност, изискват изпълнението на голям брой изчислителни операции. В някои случаи, например при генерирането на тримерни текстури, могат да се поставят ограничени изисквания за точност на апроксимацията и да се използва

само определен вид осветяване, но изискват повищено бързодействие.

Предлаганият алгоритъм за генериране на осветеност на контурни изображения използва автовълнов процес (АВП) в дискретна среда и притежава предимствата, характерни за алгоритмите, използващи АВП – високо бързодействие и възможност за паралелно изпълнение на операциите. Ограниченията на предложенията в [4] алгоритъм са във вида на градиента на яркостта (само в посока, препендикулярна на контурната крива и използването само на определен вид осветяване от направлението на наблюдателя).

Формулиране на задачата за генериране на осветяване на контурни обекти:

1. Зададено е изображение, което съдържа бял фон и обекти, ограничени със затворени контурни криви и запълнени също с бял цвят;
2. Зададена е функцията на осветяване $Fb(S)$, която изразява яркостта на точка от обекта като функция на състоянието на съответния елемент от дискретната среда;
3. Посредством автовълнов процес всеки елемент x от дискретната мрежа, който съответства на вътрешна точка на обект от изображението получава състояние $S(x)$, което е линейна функция на разстоянието на точката до контура на обекта;
4. Осветеността на точките от вътрешността на обекта се получава посредством функция на осветеност $Fb(S)$, дефинирана върху дискретното множество от възможните състояния на елементите от дискретната среда. Така посредством $Fb(S)$ се получава яркостта на точките от вътрешността на обектите в изображението.

4. Модел на АВП с N състояния на елементите

В описания в т. 2 алгоритъм за разпространение на АВП в бинарна дискретна среда се определя условието за преминаване на един елемент от състояние 0 в състояние 1 като пристап логическа функция от състоянието на съседните елементи. За генериране на осветеност контурни обекти в дискретно изображение се предлага да се използва модел на АВП с брой състояния на дискретните елементи, равен на броя на нивата на яркост, с които се кодира дискретното изображение. В използваниите графични файлови формати най-големият брой нива на яркост, с който се кодират трите основни цвята е 256 и съответно цветът на точка от изображението се задава с функцията $RGB(red, green, blue)$, където с red , $green$ и $blue$ се задава яркостта (от 0 до 255) на трите основни цвята.

В модела на АВП на точките с бял цвят съответстват пасивни елементи, на точките, а на всички останали точки (с цвят, различен от бял) съответстват активни елементи. При всяка стъпка от АВП пасивните елементи променят състоянието си само ако имат поне един съседен активен елемент. Състоянието, в което преминава пасивният елемент се определя от номера n на стъпката от изпълнение на АВП. В [1] състоянието на елемент, което не се променя при изпълнение на АВП независимо от състоянията на съседните елементи се определя като пасивно състояние. В описаният модел на АВП състоянието N (напр. $RGB(255,255,255)$) е пасивно, състоянието $N-1$ (напр. $RGB(254,254,254)$) е рефрактерно (непроменящо се), а всички останали състояния $0, \dots, N-2$ са активни.

5. Описание на алгоритъма за генериране на осветеност

Алгоритъмът включва следната последователност от операции:

1. Запълва се фонът на зададеното изображение с цвят, съответстващ на рефрактерното $N-1^{bo}$ състояние;
2. Изпълняват се $N-2$ стъпки на АВП. В съответствие с алгоритъма на АВП, на n -тата стъпка от АВП ако един дискретен елемент от изображението е в пасивно състояние и има поне една съседен елемент, който е в активно състояние, то той преминава в активно състояние n . Следователно в разглеждания модел състоянието, в което

преминава даден елемент, (когато е изпълнено условието за превключване от пасивно в активно състояние) се определя от номера на стъпката n на АВП;

3. От състоянието на елементите на дискретната мрежа посредством функцията на осветяване $Fb(S)$ се получава яркостта на точките (пикселите) от изходното изображение.

След изпълнението на стъпка 2 от алгоритъма, всеки елемент от дискретната среда, съответстващ на точка вътре в някой от контурните обекти, получава състояние, което е равно на номера на стъпката на АВП, при която вълновият процес е достигнал до елемента. Този номер в първо приближение е равен на разстоянието n от точката до контура на обекта в брой дискрети. При третата стъпка от алгоритъма всяка точка вътре в контурните обекти получава яркост Y , която е функция на разстоянието n и на вида на функцията на осветяване $Fb(S)$,

$$Y(n_1, n_2) = Fb[S(n_1, n_2)], \quad (2)$$

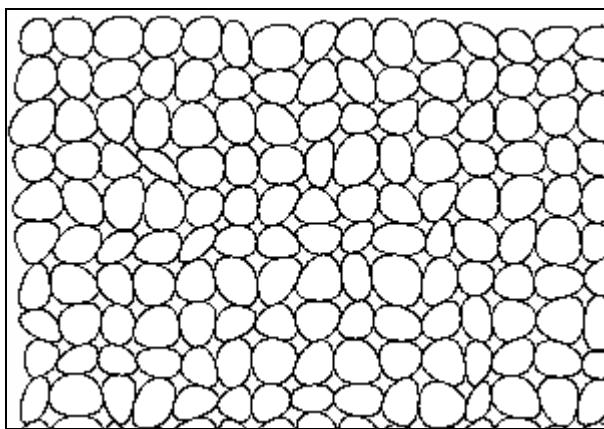
където:

$S(n_1, n_2)$ е състоянието на елемент от дискретната среда с координати n_1, n_2 ;
 $Fb(S)$ е използваната функция на осветяване.

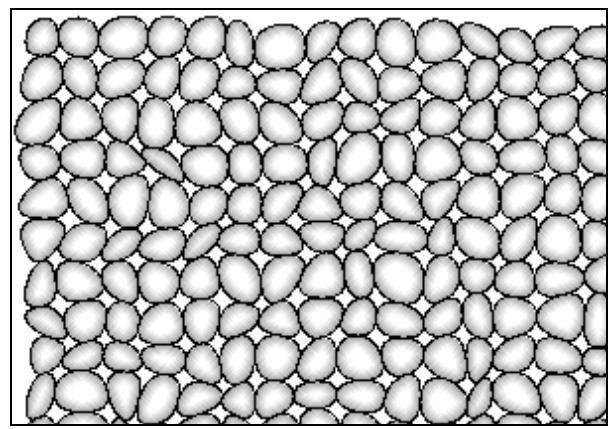
Изпълнението на стъпка 3 от алгоритъма за различни функции на осветяване ще генерира различни осветявания на контурните обекти в зададеното изображение.

6. Практически резултати

Пример за приложение на описания алгоритъм е показан на фиг. 1 и фиг. 2. Първоначалното изображение е показано на фиг. 1. То представлява текстура, получена чрез паркетиране на правоъгълник със случаен четириъгълници и последващо изглажддане на контурите на четириъгълниците.



Фиг. 1 Зададено контурно изображение



Фиг. 2 Резултат от прилагането на алгоритъма за осветяване на контурни обекти към изображението от Фиг. 1

Резултатът от изпълнението на описания алгоритъм за осветяване при 8 стъпки на АВП ($N=8$) е показан на Фиг. 2.

От резултата се вижда, че генериране на осветеност на контурните обекти в зададеното изображение създава илюзия за тримерност и значително подобрява вида на текстурата.

7. Заключение

Предлаганият алгоритъм дава възможност за генериране на осветеност на обекти в контурни изображения. Посредством изменение на яркостта на точките от вътрешността на контурните обекти, пропорционално на разстоянието от контура на обекта. Алгоритъмът притежава предимствата, характерни за автовълновите алгоритми – локалност, рекурсивност и възможност за паралелно изпълнение. За метода е характерно също така, че той не изисква предварителен анализ на формата на контурите, за които ще се генерира осветеност. Като недостатък на алгоритъма може да се посочи това, че той може да се прилага само за контурни изображения и само за светлинен източник, разположен в направлението на наблюдателя.

Литература

- [1]. В. С. Зыков, Моделирование волновых процессов в возбудимых средах. М. "Наука" 1984г.
- [2] О. Железов, Възстановяване на контурни изображения посредством автовълнов процес, списание "Автоматика и информатика", брой 5-6, 1998г
- [3] О. Железов, Генериране на осветеност на обекти посредством автовълнов процес, сборник доклади от конференция "Морски научен форум", том 2, 1998г
- [4] О. Железов, Генериране на осветеност на обекти посредством автовълнов процес, списание „Компютърни науки и технологии”, брой 1, 2006г стр. 8÷11
- [5]. Л. Рабинер, Б. Гоулд, Теория и применение цифровой обработки сигналов. М. "Мир" 1978г.
- [6]. Kung s, Whitehouse H and Kailath editors, VLSI and modern signal processing, Prentice hall inc. New Jersey, 1989
- [7]. Т. Павлидис, Алгоритмы машинной графики и обработки изображений. пер. Н.Гуревич, М. "Радио и связь", 1986г.
- [8]. D. E. Dudgeon and R. M. Mersereau, Multidimensional signal processing, Prentice Hall, Inc. 1984.

За контакти:

доц.д-р инж. Огнян Иванов Железов
катедра "Компютърни науки и технологии"
Технически университет - Варна
E-mail:ognyanz@yahoo.com
Рецензент: доц. д-р инж. Д. Генов, ТУ – Варна



МЕТОД КОНТРОЛЯ ДЛЯ ПОВЫШЕНИЯ ДОСТОВЕРНОСТИ РАБОЧЕГО ДИАГНОСТИРОВАНИЯ В ВЫЧИСЛИТЕЛЬНЫХ УСТРОЙСТВАХ

Маоэл Сайд, Александр Дрозд, Мирослав Дрозд, Геннадий Дмитриев

Резюме: Рассмотрены особенности рабочего диагностирования вычислительных устройств для выполнения приближенных вычислений. Определены области повышения достоверности рабочего диагностирования. Предложен метод контроля, который повышает достоверность, проверяя результат по упрощенной операции. Операция упрощается при вычислении результата на контрольных словах, которые выделяются из множества входных слов, используя для operandов ограничивающие условия и логические операции над ними. Метод рассмотрен на примере контроля сумматора мантисс.

Checking method to increase reliability of on-line testing in computing devices

Maoel Said, Alexander Drozd, Miroslav Drozd, Gennady Dmitriev

Abstract: Features of on-line testing in the computing devices for execution of the approximated calculations are considered. The areas of increase in reliability of on-line testing are defined. The checking method, which increases reliability, checking the result by the simplified operation is offered. The operation is simplified in case the result is calculated at the check words selected from a set of the input words using for operands the limiting conditions and logic operations under them. The method is considered on example of checking the adder of mantissas.

1. Введение

Анализ современных вычислительных устройств (ВУ) как объектов РД выделил в качестве их основной характеристики вероятность существенной ошибки (т.е., что появившаяся ошибка является существенной). Она оценивается, как $P_C = n_B / n_P$, где n_B и n_P – количество верных и вычисленных разрядов результата. В случае точных данных все ошибки являются существенными для достоверности результата, т.е. $P_C = 1$. Младшие разряды приближенных результатов являются неверными, а ошибки в них – несущественными [1].

Приближенные вычисления, как правило, выполняются в форматах с плавающей точкой [2]. Число с плавающей точкой представляется в виде произведения $m \cdot q^B$, где m – мантисса, q – основание системы счисления, B – экспонента. Присутствие умножения в самой записи приближенного числа наследует ее во все операции над мантиссами. Для полных двуместных операций произведение имеет удвоенную длину по сравнению с операндом. Причем, согласно теории ошибок, результат имеет верных разрядов не больше, чем их содержится в операнде. Поэтому для полных операций над мантиссами вероятность существенной ошибки $P_C \leq 0,5$.

Сокращенные арифметические операции вычисляют усеченный результат, обеспечивающий сохранение одинарной точности. Длина усеченного результата почти вдвое короче результата полной операции. В однотактных матричных ВУ это позволяет соответственно снизить затраты оборудования и время вычислений, а также получить близкую к единице вероятность P_C [3].

Однако, вероятность P_C дополнительно снижается для результатов всех операций, предшествующих денормализации операндов и следующих за нормализацией результата [4].

Таким образом, вероятность P_C , как правило, является низкой, и значение $P_C > 0,5$ может быть достигнуто только при выполнении сокращенных операций.

2. Оценка достоверности методов РД

Основной характеристикой методов РД является достоверность контроля результатов, которая оценивается по следующей формуле [5]:

$$D = P_C P_O + (1 - P_C)(1 - P_O), \quad (1)$$

где P_O – вероятность обнаружения ошибки.

Достоверность контроля результатов может быть представлена функцией от вероятности существенной ошибки в единичном квадрате Декартовой системы координат (рис. 1).

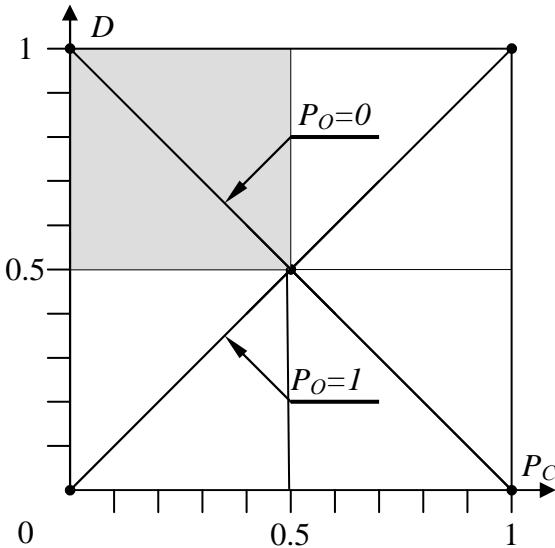


Рис. 1. Представление достоверности контроля приближенных результатов

Диагонали этого квадрата определяются отрезками прямых, задаваемых линейными функциями $D = P_C$ и $D = 1 - P_C$, которые соответствуют частным случаям формулы (1) при $P_O = 1$ и $P_O = 0$. Линии $D = 0,5$ и $P_C = 0,5$ делят квадрат на четыре области: нижние для $D < 0,5$ и верхние для $D > 0,5$.

Традиционные методы РД имеют высокую вероятность $P_O \rightarrow 1$ и для $P_C < 0,5$ занимают левую нижнюю область низких значений достоверности $D < 0,5$.

Высокая достоверность достигается для наиболее вероятного случая $P_C < 0,5$ в левой верхней области при $P_O \rightarrow 0$. Нижняя граница вероятности определяется величиной T естественной временной избыточности в форме пассивного запаса времени контроля [6], в течение которого результат остается при неисправности достоверным, как [7]

$$P_O = \ln 2 / T. \quad (2)$$

Предлагается метод контроля по упрощенной операции, повышающий достоверность контроля приближенных результатов, работая в левой верхней области.

3. Контроль приближенных результатов по упрощенной операции

Метод состоит в ограничении множества входных слов до подмножества контрольных слов (КС), на которых результат проверяется как вычисленный по упрощенной операции.

Множество КС определяется наложением ограничивающих условий (ОУ) на операнды, составляющие входные слова.

Контрольные вычисления выполняются в кодах остатков мантисс по модулю три.

Из ОУ для операндов следует ОУ для результата, используемое для его проверки.

Для операндов ОУ являются зависимыми, если определяют для результата одно и то же ОУ, и независимыми в противном случае. Независимое ОУ может быть составным и состоять из зависимых ОУ, объединенных логическими операциями. Логические операции являются допустимыми, если их искажения при неисправностях обеспечивают обнаружение ошибки, расширяя множество КС. Для зависимых ОУ допустимыми являются двуместные операции OR и XOR. Для независимых ОУ допустимой является только операция AND.

Результат оценивается как недостоверный при обнаружении ошибки на КС.

Коды остатков по модулю три принимают разрешенные значения $0I_2$ или 10_2 и запрещенные значения 00_2 или 11_2 .

Известные схемы сжатия (СС) [8], умножающие коды остатков по модулю три, выполняют операции OR и AND соответственно над разрешенными и запрещенными значениями. Инверсия одного из разрядов кода остатка переводит разрешенные значения в запрещенные или наоборот. Эту операцию NOT выполняет схема инверсии СИ. Операции OR (AND) и NOT составляют универсальный базис для реализации любой другой логической операции над ОУ, используя для обработки их кодов схемы сжатия и инверсии.

Схема контроля содержит блоки БЗ, БН и БР, формирующие зависимые и независимые ОУ для операндов и ОУ для результата, а также блок БЛ выполнения логических операций.

Контроль по упрощенной операции, показанный на примере РД однотактного сумматора мантисс, выполняет денормализацию операндов, сложение мантисс и нормализацию результата.

Исходными данными, являются n -разрядные коды мантисс операндов $A = A\{I \div n\}$ и $B = B\{I \div n\}$, величины их сдвига C_A и C_B , а также величина сдвига C_S для нормализации суммы. Вычисляются коды сдвинутых мантисс $A_C = A_C\{0 \div n\}$ и $B_C = B_C\{0 \div n\}$, их суммы $S = A_C + B_C$, $S = S\{0 \div n\}$, а также нормализованного результата $S_C = S_C\{1 \div n\}$. Пусть имеется $T = 9$ тактов пассивного запаса времени контроля. Тогда, согласно (2), вероятность обнаружения ошибки снижается до значения $P_O = 0,08$.

Денормализация операнда выполняется сдвигом его мантиссы вправо с потерей младших разрядов. Сдвинутая мантисса $A_C = A_O \cdot 2^{-CA} + A_3$, где A_O – сохраняемая часть разрядов, A_3 – высвобождаемая часть мантиссы A_C заполняемая разрядом знака.

Вводится ОУ $KA_O = 0$, которое обеспечивает для результата ОУ $KA_C = KA_3$, где $KA_C = A_C \bmod 3$, $KA_O = A_O \bmod 3$, $KA_3 = A_3 \bmod 3$. Код KA_O вычисляется, используя разбиение матрицы сдвига на фрагменты [9], как это показано для $n = 7$ на рис. 2.

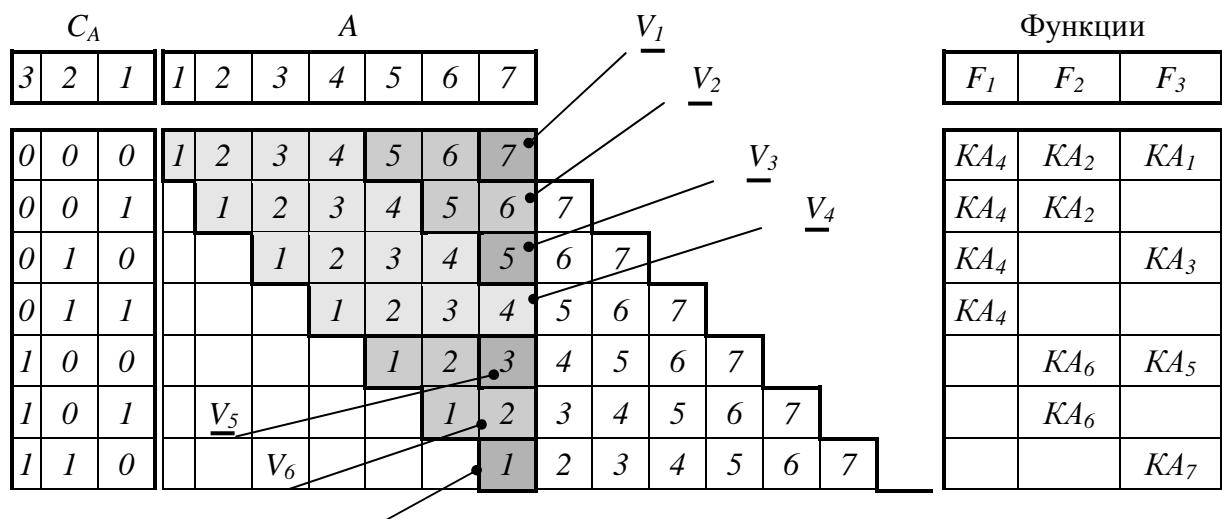




Рис. 2. Матрица сдвига

По фрагментам $V_1 \div V_7$ определяются контрольные коды по модулю три:

$$\begin{aligned} KA_1 &= -A\{7\}, & KA_2 &= A\{5, 6\}, & KA_3 &= -A\{5\}, & KA_4 &= A\{1 \div 4\} \bmod 3, \\ KA_5 &= -A\{3\}, & KA_6 &= A\{1, 2\}, & KA_7 &= -A\{1\}, \end{aligned}$$

а также функции F_1, F_2 и F_3 их выбора, реализуемые с сохранением весов разрядов мантиссы A до сдвига.

Функция F_1 выбирает KA_4 при $C_A\{3\} = 0$. Функция F_2 при $C_A\{2\} = 0$ выбирает KA_2 или KA_6 соответственно для значений 0 или 1 разряда $C_A\{3\}$. Функция F_3 при $C_A\{1\} = 0$ выбирает KA_1, KA_3, KA_5 или KA_7 соответственно для значений $00_2, 01_2, 10_2$, или 11_2 кода $C_A\{2, 3\}$.

Свертка по модулю три кодов KX_1, KX_2 и KX_3 , вычисляемых по функциям MX_1, MX_2 и MX_3 , определяет код KA_O .

Второе и третье независимые ОУ могут быть заданы как $A\{1\} = A\{2\}$ и $A\{5\} = A\{6\}$, т.е. $KA_6 = 0$ и $KA_2 = 0$. Тогда вычисляемый контрольный код преобразуется в составляемый $KA_4 = A\{3, 4\}$ и устраняется функция MX_2 , упрощая свертку по модулю три на код $KX_2 = 0$.

Аналогично выполняется контроль вычисления сдвинутой мантиссы B_C .

Контроль сдвинутых мантисс A_C и B_C при их сложении выполняется путем наложения на них ОУ $K_C = 0$, где $K_C = KA_C + KB_C$, $KA_C = A_C \bmod 3$ и $KB_C = B_C \bmod 3$, что определяет ОУ $KS = 0, KS = S \bmod 3$ для результата.

Дополнительно вводятся независимые ОУ $A_C\{n-1\} = A_C\{n\}$ и $B_C\{n-1\} = B_C\{n\}$, упрощающие вычисление кодов KA_C и KB_C на обработку используемых в ОУ разрядов.

Нормализация результата выполняется циклическим сдвигом его мантиссы S влево на величину C_A . Результат сдвига $S_C\{1 \div n\}$ определяется как $S\{0 \div n-1\}, S\{1 \div n\}$ или $S\{C_A + 1 \div n, 1 \div C_A\}$ для $C_A = 0, C_A = 1$ или $C_A > 1$, соответственно.

При $C_A \geq 1$ выполняется равенство $KA_C = (-1)^{C_A+1} KA$ и $KA = A\{0 \div n-1\} \bmod 3$, где $KA_C = A_C \bmod 3$. При $C_A = 0$ на месте знака $A\{0\}$ вычисляется разряд переполнения $A\{0\}^*$. При отбрасывании младшего разряда $A\{n\}$ равенство может нарушаться, что устраняется в случае $KA\{0\}^* = A\{n\}$, где $KA\{0\}^*$ – инверсия разряда переполнения. Для представления мантиссы A в обратном или дополнительном коде $KA\{0\}^* = -A\{0\}^* \oplus A\{0\}$. Наложение на операнд независимых ОУ $KA = 0$ и $KA\{0\}^* = A\{n\}$ определяет для результата ОУ $KA_C = 0$. Дополнительно вводится независимое ОУ $A\{n-1\} = A\{n\}$, упрощающее вычисление кода KA .

В разработанном контроле денормализации операндов, сложения сдвинутых мантисс и нормализации результата независимые ОУ (в порядке их введения) снижают мощность множества КС соответственно в 3, 2 и 2 раза, а выполнение над ними логической операции И – до величины $0,08 G$, где G – мощность множества входных слов. Это определяет требуемую вероятность обнаружения ошибки $P_O = 0,08$, для которой, согласно (1), достоверность контроля результатов составляет $D = 0,92 - 0,84 P_C$. Для сравнения, контроль по модулю три имеет $P_O = 1$ и достоверность, согласно (1), $D_M = P_C$.

На рис. 3 показаны диаграммы зависимости достоверности контроля по упрощенной операции и контроля по модулю три в однотактном сумматоре мантисс от вероятности P_C .

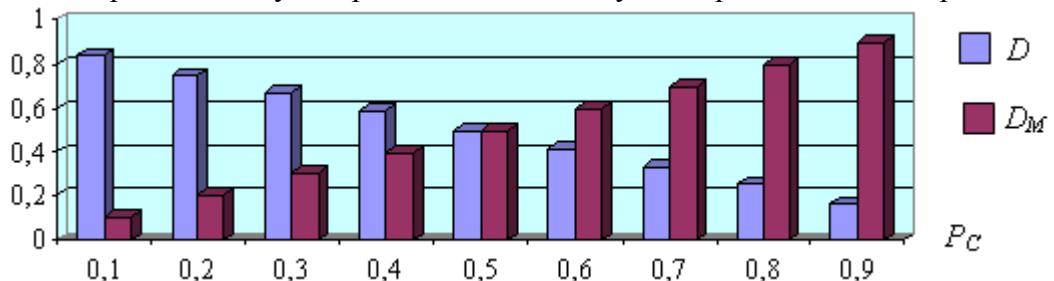


Рис. 3. Оценка достоверности контроля

Диаграммы показывают, что предложенный метод контроля однотактного сумматора мантисс дает многократный выигрыш в достоверности контроля приближенных результатов для вероятности $P_C < 0,5$ и также многократно уступает при $P_C > 0,5$. Затраты контрольного оборудования не превышают затрат оборудования в контроле по модулю три.

Предложенный контроль является целочисленным, т.е. должен строиться по самопроверяемой схеме, обнаруживающей ошибки, вызываемые собственными неисправностями. Выполнение контроля с использованием схем свертки по модулю три для формирования кодов ОУ, а также схем сжатия и инверсии для выполнения логических операций обеспечивают самопроверяемость схемы контроля в множестве одиночных константных неисправностей [10].

Предлагается расширить множество обнаруживаемых неисправностей, включая в него «закоротки» [11], возникающие между соседними разрядами кодов ОУ и результатами обработки этих кодов.

Анализ контрольных вычислений показал, что «закоротка» обнаруживается при кодировании выполнения ОУ для операндов и результата запрещенными и разрешенными значениями, соответственно. Тогда при ошибке, вызванной «закороткой» множество КС несанкционированно расширяется, а ОУ для результата нарушается, что приводит к обнаружению ошибки. При выполнении логических операций над кодами ОУ ошибка, вызванная «закороткой» не выявляется на выходах схем инверсий при переводе запрещенных значений в разрешенные.

Пусть u_1 , u_2 и z_1 , z_2 – разряды кода ОУ соответственно на входах и выходах схемы инверсии, которая вычисляет $z_1 = u_1$ и $z_2 = \neg u_2$. При «закоротке» разрядов z_1 и z_2 в зависимости от исходного кода возможны четыре случая выполнения контрольных вычислений.

В первом случае $u_1 = 0$ и $u_2 = 0$. Неисправность «закоротка» определяет $z_1 = 0$ и $z_2 = 0$ вместо $z_2 = 1$. Второй случай: $u_1 = 1$ и $u_2 = 1$. Тогда $z_1 = 0$ и $z_2 = 0$ вместо $z_1 = 1$. Третий случай: $u_1 = 0$ и $u_2 = 1$. Тогда $z_1 = 0$ и $z_2 = 0$ и это правильно. Четвертый случай: $u_1 = 1$ и $u_2 = 0$. Тогда $z_1 = 1$ и $z_2 = 1$ и это правильно. Следовательно, «закоротка» проявляется только в двух первых случаях. В первом случае характерным признаком ошибки является пара значений $u_2 = 0$ и $z_2 = 0$. Никакой другой случай не содержит такого признака, а также инверсной комбинации $u_2 = 1$ и $z_2 = 1$. Во втором случае характерным признаком ошибки является пара значений $u_1 = 1$ и $z_1 = 0$. Никакой другой случай не содержит такого признака, а также инверсной комбинации $u_1 = 0$ и $z_1 = 1$.

Для обнаружения «закоротки» в первом случае достаточно найти на схемах сжатия произведение по модулю три всех используемых в контроле пар u_2 , z_2 . Если хотя бы один из сомножителей принимает нулевое значение, то обнаруживается ошибка, вызванная «закороткой».

Обнаружение «закоротки» во втором случае выполняется по аналогии с первым случаем, если вместо разряда u_1 использовать полученную по нему инверсию $u_3 = \neg u_1$. Тогда, на схемах сжатия вычисляется произведение по модулю три всех используемых в контроле пар u_3 , z_1 . При нулевом значении сомножителя произведение принимает значение обнаружения ошибки, вызванной в контроле «закороткой».

4. Заключение

С точки зрения РД, ВУ характеризуется вероятностью существенной ошибки. Для точных данных все ошибки, вызываемые неисправностями схем, являются существенными, а их вероятность равна единице. Особенности приближенных вычислений значительно снижают вероятность существенной ошибки.

Анализ достоверности контроля приближенных результатов позволил определить области низкой и высокой достоверности методов РД.

Наибольший практический интерес представляет область наиболее вероятных низких значений вероятности существенной ошибки, где традиционные методы РД демонстрируют низкую достоверность.

Предложенный метод контроля по упрощенной операции, работающий в этой области, повышает достоверность контроля приближенных результатов, выполняя их проверку на множестве КС. Это множество определяется наложением на операнды зависимых и независимых ОУ, а также логическими операциями над ними, что устанавливает ОУ для результата. Анализ зависимых и независимых ОУ выявил допустимые логические операции, которые могут выполняться над ними с обнаружением ошибок, вызываемых одиночными константными неисправностями схем РД. Это множество неисправностей расширено включением в него также наиболее вероятных «закороток» между соседними разрядами кодов контроля.

Пример контроля однотактного сумматора мантисс показал многократное увеличение достоверности предложенного метода по сравнению с контролем по модулю три для низких значений вероятности существенной ошибки.

Литература

- [1]. А. В. Дрозд Нетрадиционный взгляд на рабочее диагностирование вычислительных устройств. Проблемы управления. № 2, 2008г.
- [2]. ANSI/IEEE Std 754-1985. IEEE Standard for Binary Floating-Point Arithmetic. – IEEE, New York, USA, 1985г.
- [3]. А. Я. Савельев. Прикладная теория цифровых автоматов. Москва Высшая школа, 1987г.
- [4]. І. Г. Мілейко, М. Б. Копитчук, О. В. Дрозд. Зниження вимог до інформаційної надійності у інформаційно-вимірювальних системах. Електромашинобудування та електрообладнання. № 72, 2009г.
- [5]. A. Drozd, M. Lobachev, J. Drozd. The problem of on-line testing methods in approximate data processing. 12th IEEE International On-Line Testing Symposium. Como, Italy, 2006г.
- [6]. А. М. Романкевич, В. Н. Валуйский, В. А. Остафин. Структурно-временная избыточность в управляющих схемах. Киев «Вища школа» 1979г.
- [7]. А. В. Дрозд Базовые концепции функционального диагностирования вычислительных устройств для приближенной обработки данных. Электронное моделирование. Том 25. № 1, 2003г.
- [8]. Е. С. Согомонян, Е. В. Слабаков. Самопроверяемые устройства и отказоустойчивые системы. Москва Радио и связь, 1989г.
- [9]. О. В. Дрозд. Контроль за модулем обчислювальних пристройв. Одеса АО Бахва, 2002г.
- [10]. Е. В. Слабаков Построение полностью самопроверяемых комбинационных схем с использованием остаточных кодов. Автоматика и телемеханика. № 10, 1979г.
- [11]. M. Blyzniuk, I. Kazymyra, W. Kuzmicz a.o. Probabilistic analysis of CMOS physical defects in VLSI circuits for test coverage improvement. Microelectronics Reliability. Vol. 41/12, 2001г.

За контакти:

проф., д-р техн. наук Александр Дрозд
кафедра компьютерных интеллектуальных систем и сетей
Одесский национальный политехнический университет – Одесса
E-mail:drozd@ukr.net
Рецензент: доц. д-р инж. С. Иванов, ТУ – Варна

Pruning methods for hidden layer of Radial Basis Function Neural Networks

Ganka P. Kovacheva

Abstract: The report tries to give a solution of the problem for determining the proper size of the hidden layer of an artificial RBF neural network by combining of fitting initialization with iteratively eliminating units from the hidden layer of the network and adjusting the remaining weights in a way that the network performance doesn't worsen over the entire training set. The idea of combining two approaches is effective, if the network contains any hidden neurons, which have almost zero weighted connections and have to be removed.

Методи за съкращаване на излишните неврони от скрития слой на невронни мрежи с радиално-базисни функции

Ганка П. Ковачева

Резюме: Докладът се опитва да даде решение на проблема за определяне на правилния размер на скрития слой на изкуствена невронна мрежа с радиално-базисни функции, чрез комбиниране на подходящо инициализиране с итеративно елиминиране на невроните от скрития слой на мрежата и настройка на останалите тегла по такъв начин, че да не се влошава поведението на мрежата над цялата обучаваща извадка. Идеята за комбиниране на двата подхода е ефективна, ако мрежата притежава скрити неврони, които имат почти нулеви тегловни връзки и трябва да бъдат отстранени.

1. Introduction

There are several algorithms that can be used to actually modify the structure of the neural network. This structural modification will not generally improve the performance of the neural network, but they can make it more efficient. This is done by analyzing how much each neuron contributes to the actual output of the neural network. If a particular neuron's connection to other neurons does not significantly affect the output of the neural network, the connection will be pruned.

Pruning [7] can be especially effective when used with a large neural network that is taking too long to execute. By pruning unneeded connections the neural network can be made to execute faster. This allows the neural network to perform more work in a given amount of time.

Reducing the number of hidden neurons in RBF layer is a very important part of deciding neural network architecture. Though this layer does not directly interact with the external environment it has a tremendous influence on the final output. Using too few neurons in the hidden layer will result in something called under – fitting, when there are too few neurons to adequately detect the signals in a complicated data set. Using too many neurons in the hidden layer can result in over - fitting. Over - fitting occurs when the neural network has so much information processing capacity that the limited amount of information contained in the training set is not enough to train all of the neurons in the hidden layer. This problem can occur even when there are sufficient training data. An inordinately large number of neurons in the hidden layer can increase the time it takes to train the network. The amount of training time can increase enough so that it is impossible to adequately train the neural network. Obviously some compromise must be reached between too many and too few neurons in the hidden layer.

There are many rule-of-thumb methods [7] for determining the correct number of neurons to use in the hidden layer. Some of them are summarized as follows.

- The number of hidden neurons should be in the range between the size of the input layer and the size of the output layer.

- The number of hidden neurons should be 2/3 of the input layer size, plus the size of the output layer.
 - The number of hidden neurons should be less than twice the input layer size.
- These three rules are only starting points that you may want to consider, but not enough to determine the network architecture.

2. Realization of Radial Basis Function (RBF) Neural Network

RBF neural networks are three (input, hidden, output) layered networks, which contain a set of locally responsive units in a hidden layer (Figure 1). Each hidden unit is a node with an RBF activation function. Every unit has two important parameters – the center and the width and shall be taken as a Gaussian function. Nodes in output layer are linear and yield a weighted sum of its inputs. The number of input nodes is equal to the dimension p of the input vector $x = (\xi_1, \xi_2, \dots, \xi_p)^T$. The number of output nodes depends on the number k of classes we want to recognize.

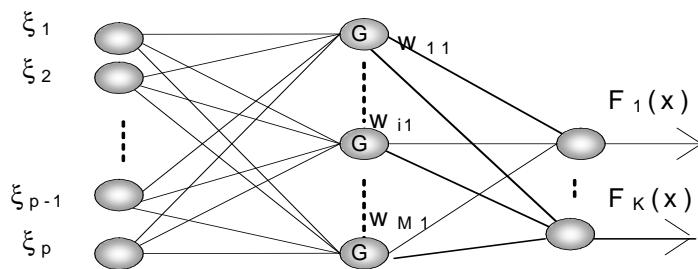


Figure 1: Radial-basis function neural network

The training data are given as a set of pairs $\{(x_i, d_i); 1 \leq i \leq N\}$, where x_i is the measurement input vector in R^p and d_i is the corresponding target vector in R^k . If x_i belongs to the class k , then d_i is the vector consisting of 0 elements except the k -th element equal to 1. Using these training data in a procedure for determining the prior domain we can determine the number M of hidden units ($M \leq N$) and initial values of their centers $\{t_i; 1 \leq i \leq M\}$ and widths $\{\sigma_i; 1 \leq i \leq M\}$.

Let $G(\xi; \sigma)$ be the Gaussian function with mean zero and the standard deviation σ [2]. The j -th output of the Gaussian RBF neural network is expressed by:

$$F_j(x) = \sum_{i=1}^M w_{ij} G(\|x - t_i\|; \sigma_i), \quad (1)$$

where $x \in R^p$ is an input vector, $\|x\|$ is the Euclidian norm of the vector x and w_{ij} is the weight parameter, which connect i -th hidden node with j -th output node.

The hidden layer's activation functions evolve slowly in accordance with some nonlinear optimization strategy, while the output layer's weights adjust themselves rapidly through a linear optimization strategy. It is necessary to involve the optimization of the size of hidden layer by using different techniques [1]:

a) Fixed centers randomly selected

The location of the centers may be chosen randomly from the training data, which are distributed in a representative manner for the problem. Specifically, a normalized RBF centered at t_i is defined as

$$G(\|x - t_i\|^2) = \exp\left(-\frac{M}{d^2} \|x - t_i\|^2\right)_{i=1,2,\dots,M}, \quad (2)$$

where M is the number of centers and d is the maximum distance between the chosen centers. The standard deviation (i.e. width) of all the Gaussian RBF is fixed at

$$\sigma = \frac{d}{\sqrt{2M}}. \quad (3)$$

Such a choice for the standard deviation σ avoids some extremes like too peaked or too flat Gaussian function but doesn't optimize the size of the hidden layer.

b) Supervised method for center selection

The centers of the radial-basis functions and all other free parameters of the network undergo a supervised learning process [5]. For a gradient-descent procedure that represents a generalization of the LMS algorithm it is necessary to define the instantaneous value of the cost function

$$\varepsilon = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^K e_{ij}^2 \quad (4)$$

where N is the number of training examples and k is the number of classes. The error signal for i -th training datum at the j -th output node is defined by

$$e_{ij} = d_{ij} - F_j(x_i) \quad 1 \leq i \leq N, 1 \leq j \leq K, \quad (5)$$

where d_{ij} is the j -th element of the target vector d_i .

It is necessary to find the parameters w_{ij} , t_i and σ_i , so as to minimize ε . Let w_i and e_j are the corresponding vectors of weights and errors. Thus, the results for weights between hidden and output layers are the following adaptive formulas [4]:

$$\frac{\partial \varepsilon(n)}{\partial w_i(n)} = \sum_{j=1}^N e_j(n) G(\|x_j - t_i(n)\| : \sigma_i(n)) \quad (6)$$

$$w_i(n+1) = w_i(n) - \eta_1 \frac{\partial \varepsilon(n)}{\partial w_i(n)} \quad i=1, 2, \dots, M \quad (7)$$

Positions of centers of the hidden units:

$$\frac{\partial \varepsilon(n)}{\partial t_i(n)} = 2w_i(n) \sum_{j=1}^N e_j(n) G'(\|x_j - t_i(n)\| : \sigma_i(n)) \sigma_i^{-2}(n) [x_j - t_i(n)] \quad (8)$$

$$t_i(n+1) = t_i(n) - \eta_2 \frac{\partial \varepsilon(n)}{\partial t_i(n)} \quad i=1, 2, \dots, M \quad (9)$$

Widths of the functions in the hidden layer:

$$\frac{\partial \varepsilon(n)}{\partial \sigma_i(n)} = -w_i(n) \sum_{j=1}^N e_j(n) G'(\|x_j - t_i(n)\| : \sigma_i(n)) Q_{ji}(n) \quad i=1, 2, \dots, M \quad (10)$$

$$Q_{ji}(n) = [x_j - t_i(n)] [x_j - t_i(n)]^T \quad (11)$$

$$\sigma_i(n+1) = \sigma_i(n) - \eta_3 \frac{\partial \varepsilon(n)}{\partial \sigma_i(n)} \quad (12)$$

The cost function ε is convex to linear parameters w_i , but non convex with respect to the centers t_i and widths σ_i and the search for the optimum values may get stuck at a local minimum in parameter space. For the initialization of the gradient-descent procedure, it is often desirable to begin the search in parameter space from a structured initial condition. It limits the region of parameter space to be searched to an already known useful area. Implementing an RBF network with prior domain selection helps to avoid falling into a local minimum and can be obtain by good initialization.

3. Initialization with prior domain selection

The design of an RBF neural network consists of determining the number of RBF functions in a hidden layer, determining their initial centers and widths and finding the weights that connect them to the output nodes. It is trivial to place RBF on every given training pattern and then make the kernel function peak when that pattern or similar patterns are presented but this becomes practically impossible, when the amount of training data is large or the learning is sequential. Sometimes when the data are not precise and even have a noise, it can cause over-fitting. Even when the number of centers is less and they are distributed in a representative manner for the problem, the learning is slowly and generalization error can not fall under definite value.

Clustering algorithm for input data is needed, like k-means algorithm or another but when a sample of one specific class is in the envelope of another class some faulty and erroneous results could be received. The clustering algorithm, which determines the number of hidden neurons, must take into account the attachment of the sample to a definite class as this would result in the accuracy and the time for training.

The clustering algorithm below [6] is appropriate for determining the number of the hidden units, their initial centers and widths. This algorithm creates a group of samples belonging to the same class, which is enough discriminate from opposite classes. The realization is made in C++ [3] and includes the following steps:

- 1) Every training point is a different cluster;
- 2) Randomly choose the label $k=1,2,3\dots c$ for every cluster;
- 3) Start from $k=1$;
- 4) Search for cluster from the same class;
- 5) Merge these two clusters and evaluate of the new centre;
- 6) Evaluate the distance d between the new centre and the centre of the nearest cluster from an opposite class;
- 7) Evaluate the distance between the new centre and the most outlying point of this cluster, which is a radius R of the cluster;
- 8) If $d>\alpha R$, where $\alpha=\text{const}$ than accept the merge from point 5 and continue from point 4, linking current k with the new created cluster and decreasing $c=c-1$.
- 9) If this inequality is not true reject the merging and restore the two initial clusters, leave k and c without changing and continue from point 4. Repeat steps from 4 to 8 for all the clusters and than increase $k=k+1$;
- 10) Repeat steps 4 to 8 until $k=c$.

Practically the range of α is $[1, 3]$ because big values cause good accuracy but limitary reduction of clusters.

The results of clustering algorithm are used for initialization of weights between input and hidden layer of the network near to a cluster centers. Thus every region of input space is localized and covered by Gaussian with calculated in advance parameters – center and width, but this is important for accuracy and not enough for network size reduction.

4. Reducing the number of neurons in the hidden layer

The individual connections between the neurons are analyzed to determine which connections have the least impact to the effectiveness of the neural network. One group of methods will remove all connections that have a weight below a certain threshold value. The second group of methods evaluates the effectiveness of the neural network as certain weights are considered for removal.

To prune individual neurons the connections between each neuron and the other neurons must be examined. If one particular neuron is surrounded entirely by weak connections to the other neurons, there is no reason to keep that neuron. To begin this process the selective pruning algorithm loops through each of the hidden neurons. For each hidden neuron encountered, the error

level of the neural network is evaluated both with and without the specified neuron. If the error rate jumps by more than a predefined level, the neuron will be left and the next neuron evaluated. If the error rate does not jump by much the neuron will be removed.

Once the program has evaluated all neurons the program repeats the process. This cycle continues until the program has made one pass through the hidden neurons without being able to remove a single neuron. Once this process is completed a new neural network is achieved that performs acceptably close to the original, yet has fewer hidden neurons. The major advantage of the selective pruning algorithm is the fact that it takes very little processing time to complete.

There is a generalized growing and pruning RBF neural network algorithm (GGAP-RBF) [8], which is based on linking the required learning accuracy with the significance of the nearest or intentionally added new neuron. Significance of a neuron gives a measure of the information content in that neuron about the function to be learned and is defined as the contribution made by that neuron to the network output averaged in the q-norm sense over all the input data received so far. GGAP is a resource – allocating network (RAN) algorithm, which means that a created network unit that consistently makes little contribution to the network's performance can be removed during the training. GGAP states a formula for computing the significance of the unit, which involves a q-fold integration of an expression using the probability density function $p(x)$ of the input data x ($x \in \mathbb{R}^q$).

In cases of sequential learning, where the data are not available all at a time as in a batch learning algorithms, a series of training samples are randomly drawn and presented to, and learned by the network one by one. GGAP for sequential learning is an approximation algorithm in which the training points are represented one by one. For an arbitrary input vector x , the network output is a linear combination of the RBF responses. If the j -th unit is removed, than the output of the network with the remaining units is a linear combination of the rest RBF functions, which results in an additional error. The significance of the j -th unit is defined as an L_q – norm of the error E , weighted by input density function $p(x)$ and for any value of q . The GGAP performs the following sequential updating of the network. A unit is initiated if significance and growth criteria are fulfilled.

The GGAP – RBF algorithm is:

1. Compute $p(x)$ based on N prehistory observations;
2. Set the desired approximation accuracy ϵ_{\min} ;
3. Set initial parameters of the network – number of units S , values of ω , μ , σ for every unit - methods for the initialization of the network are mentioned above.
4. Compute $E_{\min} = \eta \epsilon_{\min}$, where η is a user supplied parameter.
5. **FOR** each observation (x_n, y_n) presented to the network, **do**:

a) Calculate the parameters of the significance and growth criteria:

$$\begin{aligned}\varepsilon_n &= \max\{\varepsilon_{\max}, \gamma^n, \varepsilon_{\min}\}, (0 < \gamma < 1) \\ e_n &= y_n - f^{(n)}(x_n)\end{aligned}\tag{13}$$

for ε_{\min} , ε_{\max} , γ user supplied parameters and $f^{(n)}(x_n)$ the current network having S number of units;

b) Compute parameters for the next $(S+1)$ -th unit: $\omega_{S+1} = e_n$; $\mu_{S+1} = x_n$; $\sigma_{S+1} = \lambda \|x_n - \mu_{nr}\|$

c) Compute $E_{\text{sig}}(S+1)$;

d) Update the network architecture using calculations given below:

➤ **if** $\|x_n - \mu_{nr}\| > \varepsilon_n$ growth criterion and significance criterion are fulfilled then leave unit
 $S=S+1$

else

adjust the parameters of the nearest unit n_r to x_n by EKF iteration and compute: $E_{\text{sig}}(n_r)$;

➤ **if** $E_{sig}(n_r) < E_{min}$ then GGAP significance criterion is not fulfilled – remove the n_r -th unit;
endif
endif
⇒ **endfor**

5. Conclusion

The new approach of determining the proper size of the hidden layer of RBF neural network consists of combining an initialization with prior domain selection and iterative elimination of units from the hidden layer of the network and adjusting the remaining weights in a way that the network performance doesn't worsen over the entire training set. This approach is adaptive to arbitrary training sets, especially if the training data contain many fluctuations. An experimental study of the GGAP-RBF after initialization by clustering algorithm is conducted, following the established rules for normalization of the attributes of the data sets into the range [0,1] and the values of the GGAP user-supplied parameters, which have been set using the guidelines: $\varepsilon_{max}=1.15$, $\varepsilon_{min}=0.04$, $\gamma=0.999$ and $\lambda=0.1$ and $0.03 < \eta < 0.1$. The GGAP-RBF is a sequential algorithm and can be used for on-line learning in real time applications, where the training observations are presented one-by-one and discarded after being learned and the parameter adjustment and pruning are conducted whenever a new observation is presented.

Literature

- [1]. Andrews R. and Geva S.: "On the effects of initializing a neural network with prior knowledge", ICONIP'99, Perth, Australia, vol.1, pp.251-256, 1999
- [2]. Haykin S.: "Neural networks", Mc Master University, 1994
- [3]. Kovacheva G. and Ogawa H.: "Radial basis function classifier for fault diagnostics", ISICT'03, Dublin, Ireland, 2003
- [4]. Kovacheva G. and Ogawa H.: "Incremental learning method for RBF classifiers", WISICT'04, Cancun, Mexico, 2004
- [5]. Moody J. and Darken C.: "Fast learning in network of locally- tuned processing units", Neural Computation, 1, 1989
- [6]. Musavi M., Ahmed W. and Chan K.: "On the training of Radial Basis Function classifiers", Neural Networks, 5, pp. 595-603, 1992
- [7]. Guang - Bin Huang, Saratchandran P., Sundararajan N.: A generalized growing and pruning RBF neural network for function approximation, Neural Networks, 16, pp. 57-67, 2005
- [8]. Bortman M. and M. Aladjem: A growing and pruning method for RBF networks, Neural Networks, 20, pp. 1039-1045, 2009

За контакти:

гл. ас. д-р инж. Ганка Петкова Ковачева
катедра „Компютърни науки и технологии”
Технически университет – Варна
E-mail: gpp_k@mail
Рецензент: доц. д-р инж. В. Наумов, ТУ – Варна

Обобщенная модель беспроводных каналов стандарта ieee 802.11

Сергей А. Нестеренко, Сергей С. Нестеренко, Юлия С. Нестеренко

Резюме: Предложена обобщенная аналитическая модель беспроводных каналов стандарта IEEE 802.11. Модель учитывает коллизии и уровень битовых ошибок в беспроводном канале связи. Экспериментальным методом проведена оценка погрешности аналитической модели.

The Generalized Model of Wireless Channels Standard IEEE 802.11

Sergej A. Nesterenko, Sergej S. Nesterenko, Ulija S. Nesterenko

Abstract: The generalized analytical model of the wireless channels standard IEEE 802.11 is offered. The model takes into account collisions and level of bit errors in the wireless communication channel. The experimental method is used for an error estimation of analytical model.

1. Введение

Характерной особенностью развития современных компьютерных сетей является использование в их составе зон беспроводного доступа. Стремительное развитие беспроводных технологий и уменьшение стоимости соответствующего коммуникационного оборудования обеспечивает постоянный рост количества беспроводных абонентов корпоративных компьютерных сетей. Качественное проведение проектных работ по вводу в эксплуатацию беспроводных сегментов корпоративной сети требует наличия соответствующих моделей расчета их производительности. Работа физического и канального уровней беспроводной сети определена группой стандартов IEEE 802.11 [1]. Модификации стандарта отличаются реализацией на физическом уровне, но используют единый алгоритм работы на канальном уровне CSMA/CA (множественный доступ с прослушиванием несущей и избежанием коллизий). Использование единого метода доступа позволяет построить обобщенную информационную и математическую модель для всех модификаций стандарта.

В качестве скоростных характеристик беспроводного канала в стандарте приводится скорость передачи информации на физическом уровне. Для пользователя существенно более важной характеристикой является скорость передачи на канальном уровне, которая учитывает задержки, вносимые алгоритмом работы протокола, коллизии, возникающие при передаче кадров, а также повторные передачи, связанные с определенным уровнем помех в канале.

В данной статье проводится разработка обобщенной аналитической модели расчета пропускной способности моноканала для различных реализаций беспроводной сети стандарта 802.11 с учётом коллизий и уровня битовых ошибок в канале связи (BER – bit error rate).

2. Принципы работы беспроводных сетей стандарта IEEE 802.11

Моноканал беспроводной сети стандарта IEEE 802.11 представляет собой среду, разделяемую беспроводными устройствами: абонентами и точкой доступа. В сети инфраструктурой точки доступа – центральный элемент, через него происходит обмен данными между компонентами беспроводной сети, а также компонентами проводной и беспроводной сетей. Метод доступа к среде CSMA/CA работает следующим образом.

Станция, желающая передавать, тестирует канал, и если не обнаружено активности в течение интервала времени DIFS (DCF interframe space), ожидает случайный промежуток времени back off period, а затем передаёт, если канал всё ещё свободен. Если кадр приходит целым, принимающая станция ожидает интервал времени SIFS (short interframe space) и затем посыпает кадр подтверждения ACK (acknowledge). Если передающая станция не получила кадр ACK, делается предположение, что произошла коллизия, и кадр данных передаётся снова через случайный промежуток времени. Интервал back off period состоит из определённого количества интервалов времени – слотов. Количество слотов – равномерно распределённая случайная величина. Если несколько станций соревнуются за доступ к моноканалу и у одной из станций back off period заканчивается, она начинает передавать, а другие станции переводятся в режим ожидания. В течение интервала back off в каждом слоте станции проверяют, свободна ли среда, и если среда свободна, уменьшают значения счётчика back off, в противном случае – приостанавливают процедуру и замораживают текущие значения счётчиков. Когда среда освобождается вновь, эти станции соревнуются с уже имеющимися значениями back off.

Канальный уровень сетей стандарта 802.11 определяет два режима передачи: с предварительной установкой соединения и без неё. Режим с предварительной установкой соединения использует метод Request to Send/Clear to Send (RTS/CTS), в котором посылающая станция передаёт кадр RTS и ждёт ответа точки доступа с кадром CTS. Сигнал CTS заставляет все станции в сети отложить свои передачи на время завершения сеанса связи и получения кадра ACK передающей станцией. Использование метода RTS/CTS опционально и задаётся на стадии инициализации сети.

3. Построение аналитической модели расчёта пропускной способности моноканала

С учетом концептуального описания работы моноканала, приведенного выше, полное время передачи кадра при использовании механизма RTS/CTS вычисляется по следующей формуле:

$$T = t_D + t_R + 3 \cdot t_S + t_C + t_I + t_A + t_W, \quad (1)$$

где t_R – время передачи кадра RTS,

t_C – время передачи кадра CTS,

t_I – время передачи кадра данных,

t_A – время передачи кадра ACK,

t_D – время ожидания интервала DIFS,

t_S – время ожидания интервала SIFS,

t_W – время ожидания случайного интервала back off period.

Время передачи кадра без использования механизма RTS/CTS определяется в виде:

$$T = t_D + t_I + t_S + t_A + t_W. \quad (2)$$

Время передачи кадра RTS, CTS и ACK рассчитывается по следующей формуле

$$t_M = \frac{L}{V_S} + t_P,$$

где V_S – скорость передачи данных, определяемая версией стандарта IEEE 802.11, L – длина кадра в байтах,

t_p – время передачи преамбулы и заголовка PLCP кадра (physical layer convergence protocol – протокол сближения физического уровня) [2].

Время передачи данных, определяемое формулами (1) и (2) необходимо скорректировать добавлением времени, связанного с возникновением и обработкой коллизии.

Время обработки коллизии состоит из времени повторной передачи и времени ожидания тайм-аута ACKTimeout протокола канального уровня. В работе [3] получено аналитическое выражение для расчета вероятности коллизий P_c в зависимости от режима работы протокола R , числа абонентов N и размера кадра S . Время передачи данных с учётом коллизий T_C может быть рассчитано в виде:

$$T_C = T + P_C \cdot (t_{AT} + t_W)$$

где t_{AT} – длительность таймаута ACKTimeout.

Длительность таймаута ACKTimeout без использования режима работы RTS/CTS вычисляется по следующей формуле:

$$t_{AT} = t_I + 2 \cdot t_S + t_A$$

Длительность таймаута ACKTimeout при использовании режима работы RTS/CTS определяется в виде:

$$t_{AT} = t_R + 2 \cdot t_S + t_C$$

Пропускная способность беспроводного моноканала V для заданного размера кадра S вычисляется следующим образом:

$$V(S) = \frac{1}{T_C} \cdot S$$

В работе [4] показано, что вероятность повторной передачи кадра при уровне битовых ошибок в канале равном P_{OB} вычисляется в виде $P_{PP} = \frac{SP_{OB}}{1 - SP_{OB}}$. Повторная передача осуществляется после истечения величины тайм-аута T_{TA} протокола транспортного уровня TCP. Следовательно, время транзакции кадра в беспроводном моноканале T_{TP} стандарта 802.11 с учетом коллизий и битовых ошибок можно представить в виде

$$T_{TP} = T + P_C (t_{AT} + t_W) + \frac{T_{TA} SP_{OB}}{1 - SP_{OB}}. \quad (3)$$

Пропускная способность беспроводного канала связи $V(S)$ при передаче кадров размера S вычисляется по формуле

$$V(S) = \left[T + P_C (t_{AT} + t_W) + \frac{T_{TA} SP_{OB}}{1 - SP_{OB}} \right]^{-1} \cdot S. \quad (4)$$

Очевидно, что выражения (3,4) являются обобщенной аналитической моделью беспроводного канала стандарта 802.11, работающего в условиях коллизий и битовых помех.

Экспериментальным образом для версии канала IEEE 802.11b была проведена оценка погрешности обобщенной аналитической модели беспроводного канала. Оценивалась относительная погрешность расчета пропускной способности канала в виде

$$R = \frac{|V_{BK}^{\mathcal{E}} - V_{BK}^A|}{V_{BK}^{\mathcal{E}}},$$

где $V_{BK}^{\mathcal{E}}$ – экспериментально полученные данные скорости передачи в беспроводном канале, V_{BK}^A – данные полученные в результате аналитического моделирования. Результаты расчета погрешности модели R для различных режимов работы беспроводного канала приведены на рис. 1.

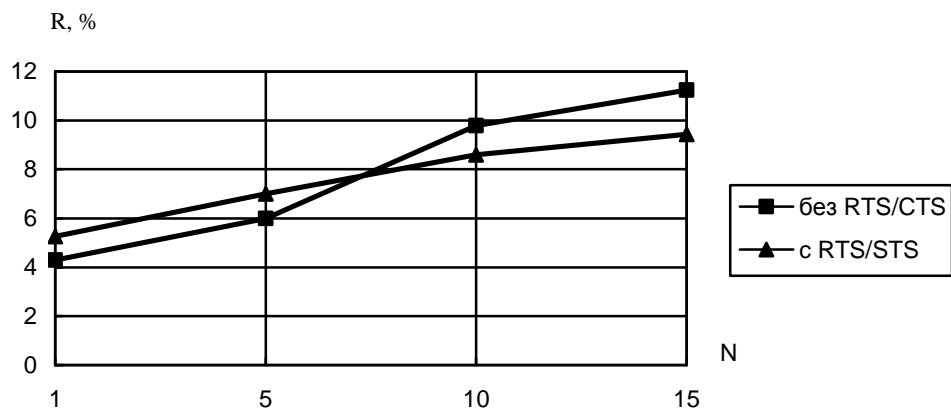


Рис. 1. Расчет погрешности аналитической модели для различных режимов работы беспроводного канала связи

С использованием аналитической модели расчета пропускной способности беспроводного канала связи проведено исследование влияния битовых ошибок на производительность канала связи для различных уровней битовых ошибок. Результаты исследования приведены на рисунке 2. Анализ результатов показывает, что характеристики беспроводного моноканала существенно ухудшаются при уровне битовых ошибок 10^{-5} . При повышении уровня битовых ошибок до величины 10^{-4} пропускная способность беспроводного канала уменьшается до величины менее 0,5 Мбит/с, что делает использование канала мало эффективным.

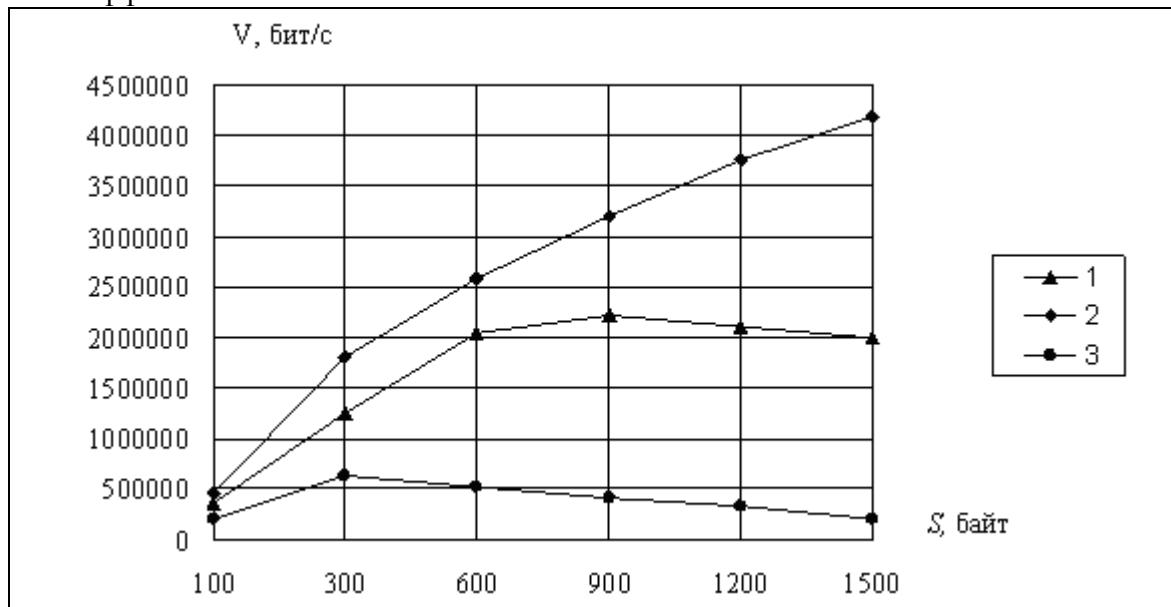


Рис. 2. Пропускная способность реального беспроводного моноканала при различных уровнях битовых ошибок в канале связи:

1 - уровень ошибок 10^{-6} ; 2 - уровень ошибок 10^{-5} ; 3 - уровень ошибок 10^{-4} .

4. Выводы

Анализ погрешности обобщенной аналитической модели показывает, что она не превышает 12% для всех режимов работы беспроводного канала, что свидетельствует о ее

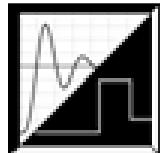
высокой точности и позволяет использовать в задачах анализа и синтеза беспроводных компьютерных сетей, которые проектируются на базе всех модификаций стандарта 802.11

Литература

- [1]. Столингс Вильям. Беспроводные линии связи и сети. Москва, Издательский дом «Вильямс», 2003, 640 с.
- [2]. ANSI/IEEE Standard 802.11, 1999 Edition – <http://standards.ieee.org/802.11-1999.pdf>
- [3]. Нестеренко С.А. А.Ю. Биньковский, Расчет полезной пропускной способности моноканала беспроводной сети стандарта 802.11 с учетом коллизий. Одесса, Холодильна техніка і технологія, 2005, № 6 (98), С. 125-128.
- [4]. Нестеренко С.А., А.Ю. Биньковский. Метод расчета пропускной способности моноканала беспроводной сети стандарта 802.11. Одесса, Праці УНДІРТ, 2006, № 1 (45), С. 109-113.

Для контактов:

д.т.н., проф. Нестеренко Сергей Анатольевич,
Одесский национальный политехнический университет, проректор,
e-mail: sa_nesterenko@ukr.net,
Адрес: пр. Шевченко, 1,
Одесский национальный политехнический университет,
Одесса 65001, Украина
Рецензент: доц. д-р инж. Й. Урумов, ТУ – Варна



ЛОКАЛИЗАЦИЯ КВАЗИПЕРИОДИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ НА ИЗОБРАЖЕНИЯХ

Светлана Г. Антощук, Анатолий А. Николенко, Тиен Т. К. Нгуен

Резюме: Предложена методика локализации квазипериодических последовательностей на изображении на основе учета свойств интегрального распределения энергии по масштабам непрерывного вейвлет-преобразования для базисных функций Гаусса и Хаара. Показано, что интегральное распределение энергии по масштабам непрерывного вейвлет-преобразования для модели строки объекта и модели строки символьной последовательности имеет характерные особенности, которые могут быть использованы при выборе рационального значения масштабного уровня для локализации в области вейвлет- преобразования отдельных объектов либо группы объектов на изображении. Результаты моделирования подтвердили работоспособность предложенной методики.

Localization of Quasi Periodic Sequences at the Images

Svetlana G. Antoshchuk, Anatoly A. Nikolenko, Tien T. K. Nguyen

Abstract: Method of localization of quasi periodic sequences at the image adjusted for integral energy distribution on continuous wavelet transform scales for the base functions of Gauss and Haar is offered. It is showed that integral energy distribution on continuous wavelet transform scales for the object row model and character sequence row model has characteristic features which can be used for the choice of rational value of scale level for localization in area of wavelet transform of separate objects or group of objects at the image. Simulations results confirmed method of localization efficiency.

1. Введение

Одной из базовых процедур при распознавании объектов на изображении является их локализация, определяющая эффективность, оперативность и ресурсоемкость автоматизированной системы обработки и распознавания изображений в целом. Локализация объектов чаще всего основывается на контурной сегментации, позволяющей существенно сократить объем обрабатываемой информации и обеспечить инвариантность к трансформациям интенсивности. Контурная сегментация, как правило, состоит из двух этапов: подчеркивание перепадов интенсивности и принятие решения о наличии границы объекта.

Для реализации первого этапа при решении ряда практических задач, связанных с локализацией объектов, которые могут быть представлены квазипериодическими последовательностями (тексты, штриховые коды, периодические текстуры и т.д.), применяется непрерывное вейвлет-преобразование (ВП) [1,2,3,4], позволяющее подчеркнуть и расширить границы объектов, изменяя степень детализации. Однако вейвлет-анализ оставляет большую свободу в конкретном выборе базисных функций с учетом характера изображения и решаемой задачи. В статье предложена методика локализации квазипериодических последовательностей, базирующаяся на обоснованном выборе масштаба ВП с учетом характера интегрального распределения энергии по масштабным уровням.

2. Общие требования к вейвлет-функциям

Для локализации квазипериодических последовательностей на изображении, как правило, применяют построчную (либо постолбцовую) обработку (функция интенсивности в строке (столбце) рассматривается как одномерный сигнал), с учетом пространственно-частотных характеристик последовательности. Для этого применяют действительные вейвлеты, заданные в виде симметричных нечетных функций $\psi(t) = -\psi(-t)$, имеющие компактный или эффективный носитель, в частности, вейвлеты Гаусса, Хаара, гиперболическое вейвлет-преобразование (ГВП) и т.д. [2, 4]. Эти базисные вейвлет-функции (ВФ) удовлетворяют ряду необходимых требований: локализации, допустимости, осцилляции (знакопеременности), ограниченности

Следует отметить, что несимметричные ВФ не используются для подчеркивания перепадов интенсивности изображения, т.к. смещают центр перепада интенсивности. В тоже время, симметричные ВФ обладают свойством сохранять неизменным положение центра перепада при его подчеркивании. Благодаря этому качеству они находят применение при локализации объектов, основанной на методах обнаружения перепадов интенсивности на изображении.

Непрерывное ВП (НВП) с действительной ВФ определяется выражением [2]

$$WF(x, s) = \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-x}{s}\right) dt, \quad (1)$$

где $f(t)$ — преобразуемая (анализируемая) функция; $\psi(t, s) = \psi\left(\frac{t-x}{s}\right)$ — базисная функция (вейвлет) с единичной нормой для всех масштабов s ; x — параметр сдвига.

3. Интегральное распределение энергии по масштабам вейвлет-преобразования

Графическому образу строки с квазипериодической последовательностью соответствует модель квазипериодического сигнала в виде прямоугольных импульсов [4]

$$f(t) = \sum_{i=1}^{i=N} 1\left(t - x_0 - (i-1)\Gamma - \sum_{k=2}^i \varepsilon_{Tk}\right) - \sum_{i=1}^{i=N} 1\left(t - x_0 - (i-1)\Gamma - \sum_{k=2}^i \varepsilon_{Tk} - d - \xi_i\right), \quad (2)$$

где N — количество прямоугольных импульсов последовательности; $1(t)$ — единичная функция Хевисайда; x_0 — координата первого импульса; ε_{Tk} и ξ_i — случайные изменения периода Γ и длительности d импульса соответственно.

Для упрощения выкладок рассмотрен частный случай модели (2) при $\varepsilon_{Tk} = 0$, $\xi_i = 0$ (т.е. период Γ и длительность d постоянны).

Модель строки, содержащей однородный объект представляется одним импульсом ($N = 1$) с длительностью, равной протяженности объекта $d = D$, а для модели строки с символьной последовательностью или с периодической текстурой $N > 1$.

Найдем функцию интегрального распределения энергии (ИРЭМ) по масштабным уровням НВП квазипериодических сигналов $f(t)$ [3]

$$En(s) = \int_{-\infty}^{\infty} |WF(x, s)|^2 dx, \quad (3)$$

где $WF(x, s)$ – НВП сигнала $f(t)$.

Поскольку аналитические выкладки в пространственной области слишком громоздки, то анализ НВП $WF(x, s)$ проведем в частотной области, используя Фурье-образ $\hat{WF}(\omega, s)$.

С учетом свойств преобразования Фурье выражение (1) имеет вид:

$$\hat{WF}(\omega, s) = \Phi(f(t)) \cdot \Phi(\psi(t, s)) = \hat{F}(\omega) \cdot \hat{\psi}(\omega, s), \quad (4)$$

где Φ – оператор преобразования Фурье, $\hat{F}(\omega)$, $\hat{\psi}(\omega, s)$ – Фурье-образы $f(t)$ и $\psi(t, s)$ соответственно.

Функция ИРЭМ (3) в частотной области определяется как

$$En(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{WF}(\omega, s)|^2 d\omega \quad (5)$$

Преобразования Фурье НВП модели (2) имеет вид

$$\hat{WF}(\omega, s) = \frac{2}{\omega} \exp\left(-j\omega\left(x_0 + \frac{d}{2}\right)\right) \times \left(\sum_{i=1}^N \sin\left(\omega \frac{d + \xi_i}{2}\right) \exp\left(-j\omega\left[(i-1)T + \sum_{k=1}^i \varepsilon_{T_k} + \frac{\xi_i}{2}\right]\right) \right) \hat{\psi}(\omega, s). \quad (6)$$

С учетом четности подынтегрального выражения (6) получим

$$En(s) = \frac{4}{\pi} \int_0^{\infty} |\hat{\psi}(\omega, s)|^2 \cdot \frac{\sin^2 \frac{\omega d}{2} \sin^2 \frac{N\omega T}{2}}{\omega^2 \sin^2 \frac{\omega T}{2}} d\omega. \quad (7)$$

Проведем расчет и анализ ИРЭМ для базисных ВФ Хаара (HAAR-вейвлет) и Гаусса (WAVE-вейвлет). Аналитическое представление этих функций и их ИРЭМ представлены в табл. 1. Эти ВФ обладают различными фильтрующими свойствами. ВФ Гаусса имеет гладкую передаточную характеристику и наилучшие фильтрующие свойства, что приводит к «размытию» границ объектов. ВФ Хаара имеют боковые лепестки в передаточных характеристиках, что обеспечивает возможность подчеркивания более высокочастотных составляющих с меньшей интенсивностью.

Таблица 1 Базисные функции и ИРЭМ квазипериодической последовательности

Функция	Аналитическое представление	ИРЭМ $En(s)$	Значение предела $E_{\infty}(s)$
Хаара	$\psi(t, s) = \begin{cases} -\frac{1}{\sqrt{s}}, & -0.5 \leq t < 0; \\ +\frac{1}{\sqrt{s}}, & 0 \leq t \leq 0.5 \end{cases}$	$\frac{64}{\pi s} \int_0^{\infty} \frac{\sin^4 \frac{\omega s}{4} \sin^2 \frac{\omega d}{2} \sin^2 \frac{N\omega T}{2}}{\omega^4 \sin^2 \frac{\omega T}{2}} d\omega$	$N^2 d^2$

Гаусса	$\psi(t, s) = \frac{1}{\sqrt[4]{2\pi}} \frac{t}{s\sqrt{s}} \exp\left(-\frac{t^2}{2s^2}\right)$	$\frac{4\sqrt{2}s^3}{\sqrt{\pi}} \int_0^\infty \exp(-\omega^2 s^2) \frac{\sin^2 \frac{\omega d}{2} \sin^2 \frac{N\omega T}{2}}{\sin^2 \frac{\omega T}{2}} d\omega$	$\frac{1}{2\sqrt{2}} N^2 d^2$
--------	--	--	-------------------------------

Анализ показал, что функция ИРЭМ НВП с базисной ВФ Хаара и Гаусса при возрастании значения масштаба s стремится к пределу, который равен $E_\infty(s) = N^2 d^2$ и $E_\infty(s) = \frac{1}{2\sqrt{2}} N^2 d^2$ соответственно.

4. Моделирование и анализ полученных результатов

Проведено компьютерное моделирование с использованием численных методов при расчете ИРЭМ (рис. 1). Моделирование проводилось для модели строки с однородным объектом ($N = 1$) и модели строки с символами ($N = 19$).

Результаты моделирования показали, что для строки изображения, содержащей однородный объект, с ростом масштабного уровня s НВП (как с базисной ВФ Гаусса, так и с базисной ВФ Хаара) наблюдается увеличение эффекта подчеркивания границ (рис. 1, а, б). Это соответствует росту ИРЭМ (кривая 1) с увеличением масштабного уровня s (рис. 1, д, е), т.е. эта функция ИРЭМ для однородного объекта имеет монотонно возрастающий характер и стремится к пределу, пропорционально зависящему от длительности импульса d .

Анализ НВП строки символов показывает, что на низких масштабных уровнях подчеркиваются перепады интенсивности на границах отдельных символов, а на высоких – перепады на границе всей надписи в строке (рис. 1, ж, з). Полученная зависимость ИРЭМ для модели строки символов имеет экстремальный характер (рис. 1, д, е) – кривая 2.

Максимальному значению ИРЭМ (рис. 1, д, е) соответствует масштабный уровень s_1 , обеспечивающий наилучшее подчеркивание отдельных символов надписи, минимальному значению – масштабный уровень s_2 НВП, после которого строка символов воспринимается как однородный объект, что соответствует положению границ надписи.

Следует отметить, что вид НВП периодической последовательности с базисной ВФ Гаусса и с базисной ВФ Хаара отличаются (рис. 1, в, г). Поскольку функция Гаусса имеет гладкую передаточную характеристику, то с ростом масштабного уровня НВП наблюдается увеличение эффекта подчеркивания внешних границ последовательности. ИРЭМ имеет две экстремальные точки – кривая 2 (рис. 1, д, е). Значение масштабного уровня, соответствующее локальному максимуму, характеризуется резонансом (совпадением периодов) ВФ и последовательности импульсов и позволяет зафиксировать необходимый масштабный уровень (s_1) для обнаружения отдельных элементов данной последовательности. При достаточно большом масштабном уровне НВП s_2 (превышающем значение, соответствующее последнему локальному минимуму) квазипериодическая последовательность воспринимается как однородный объект, отдельные импульсы последовательности становятся неразличимы.

В базисных функциях Хаара имеются боковые лепестки в передаточных характеристиках, что обеспечивает возможность подчеркивания более высокочастотных составляющих с меньшей интенсивностью. НВП носит пульсирующий характер (рис. 1, г). Функция ИРЭМ имеет дополнительно несколько перегибов или менее ярко выраженных экстремумов. Все последующие максимумы повторяются через интервал, равный удвоенному периоду следования импульсов (рис. 1, е) – кривая 2. Следовательно, по количеству локальных экстремумов можно оценить количество импульсов в последовательности. При достаточно большом масштабном уровне s НВП (превышающем

значение, соответствующее последнему локальному минимуму) квазипериодическая последовательность воспринимается так же, как и при базисной функции Гаусса, как однородный объект, отдельные импульсы последовательности становятся неразличимы. Поэтому для обнаружения границ последовательности можно использовать значение масштабного уровня, равное величине, при которой наблюдается последняя точка перегиба зависимости ИРЭМ НВП. При дальнейшем увеличении масштабного

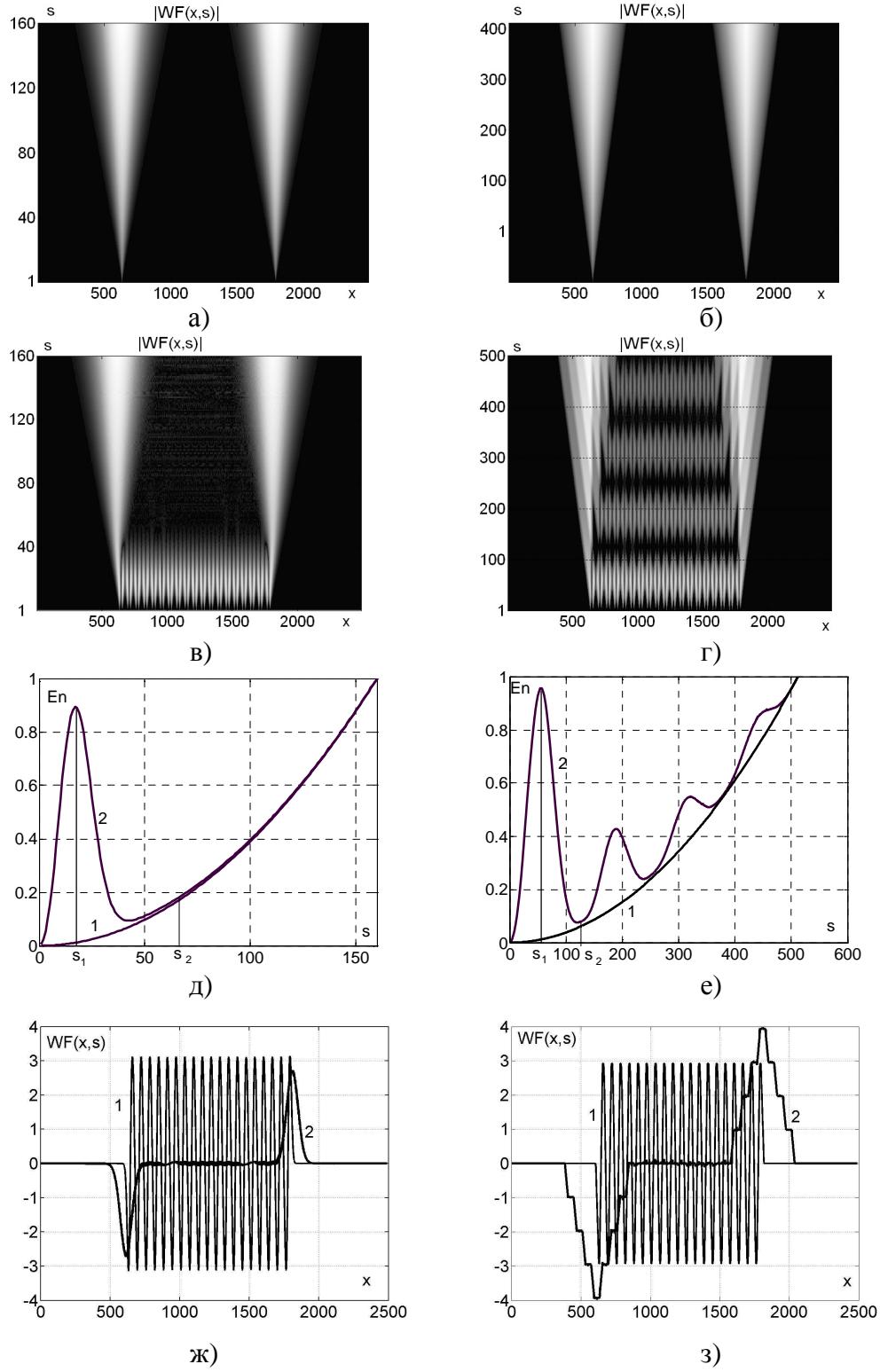


Рис. 1. Результаты моделирования:

НВП для модели строки с объектом (одиночным импульсом) и символьной строкой при ВФ Гаусса (а, в) и ВФ Хаара (б и г);
зависимость ИРЭМ для моделей строки с объектом (1) и символьной строкой (2): при ВФ Гаусса (д) и ВФ Хаара (е);

коэффициенты ВП на разных масштабах $s_1(1) \ll s_2(2)$: ВФ Гаусса (ж), ВФ Хаара (з).
уровня НВП вид функции ИРЭМ становится монотонно возрастающим и стремится к некоторому предельному значению, определяемому типом базисной функции и количеством импульсов в квазипериодической последовательности.

Исследование зависимости ИРЭМ НВП от масштабного уровня для разных базисных функций показало, что многоэкстремальный характер этой зависимости наблюдается только при наличии значительных боковых лепестков в их АЧХ (вейвлет Хаара, гиперболическое вейвлет-преобразование). Для «гладких» вейвлет-функций многоэкстремальность не наблюдается.

На основе установленной зависимости ИРЭМ предложена методика локализации квазипериодических последовательностей с учетом обоснованного выбора масштабного уровня НВП, которая заключается в следующем:

1. Проводится НВП исходного изображения по строкам (либо столбцам).
2. Вычисляется функция ИРЭМ НВП и проводится ее анализ.
3. При наличии экстремумов в функции ИРЭМ определяются масштабы для обработки периодической последовательности:
 - по положению первого максимума ИРЭМ определяется требуемый масштабный уровень НВП для подчеркивания каждого символа последовательности;
 - по положению последней точки перегиба ИРЭМ определяется требуемый масштабный уровень НВП для локализации всей последовательности;
4. Осуществляется локализация квазипериодической последовательности путем подчеркивания ее границ на выбранном масштабе.

Предложенная методика была проверена на тестовых примерах изображений и показала приемлемый результат.

5. Заключение

Таким образом, при локализации квазипериодических последовательностей на изображении целесообразно использовать интегральное распределение энергии по масштабам НВП. Его анализ позволяет получить информацию о структуре изображения (например, наличие строки символов или однородного объекта) и обеспечить выбор требуемого масштабного уровня НВП. Разработанная методика локализации квазипериодических последовательностей представляет интерес для специалистов, занимающихся обработкой сигналов и изображений в прикладных системах самого широкого применения.

Литература

- [1]. Гонсалес Р., Р. Вудс. Цифровая обработка изображений. М. , Техносфера, 2005.
- [2]. Малла С. Вейвлеты в обработке сигналов. М., Мир, 2005.
- [3]. Короновский А. А., А. Е. Храмов. Непрерывный вейвлетный анализ и его приложения М., Физматлит, 2003.
- [4]. Антощук С. Г. , А. А. Николенко, Е. В. Ткаченко, О. Ю. Бабилунга. Использование вейвлет-преобразования при локализации последовательностей символов. Искусственный интеллект. 2009, № 4, с. 23 - 29.

Для контактов:

проф. д-р техн. наук Светлана Антощук
кафедра Информационных систем
Институт компьютерных систем
Одесский национальный политехнический университет
E-mail:asg@ics.opu.ua

Рецензент: проф. д.т.н. инж. К. Тенекеджиев, ТУ – Варна



Ключови генератори за потокови шифри на базата на преместващи регистри с линейни обратни връзки

Петър Ц. Антонов

Резюме: Предлагат се схеми на генератори на ключови потоци, реализирани на базата на преместващи регистри с променливи линейни обратни връзки. Схемите са обобщени и могат да бъдат конкретизирани и използвани за потоково шифриране.

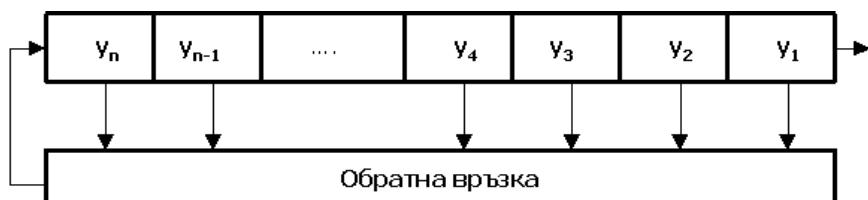
Key Generators for Stream Ciphers on the Basis of Linear Feedback Shift Registers

Peter. Ts. Antonov

Abstract: Schemes of keystream generators based on variable linear feedback shift registers are offered. Schemes are universal and they can be specified and used for stream ciphering.

Въведение

Както е известно, потоковите криптографски алгоритми (шифри) се използват широко в практиката на криптографската защита [1,2,4,6 и др.]. При това, в много случаи необходимите за целта генератори на ключови потоци се реализират аппаратно на основата на така наречените **преместващи регистри с обратни връзки** - ПРОВ (FSR - Feedback Shift Registers). Обобщената блок схема на такъв n -разряден регистър е представена на фиг. 1, където обратната връзка може да бъде **линейна** или **нелинейна**. В зависимост от това, преместващите регистри са с **линейни** (LFSR - Linear FSR) или **нелинейни** (NLFSR - Nonlinear FSR) обратни връзки.



Фиг. 1

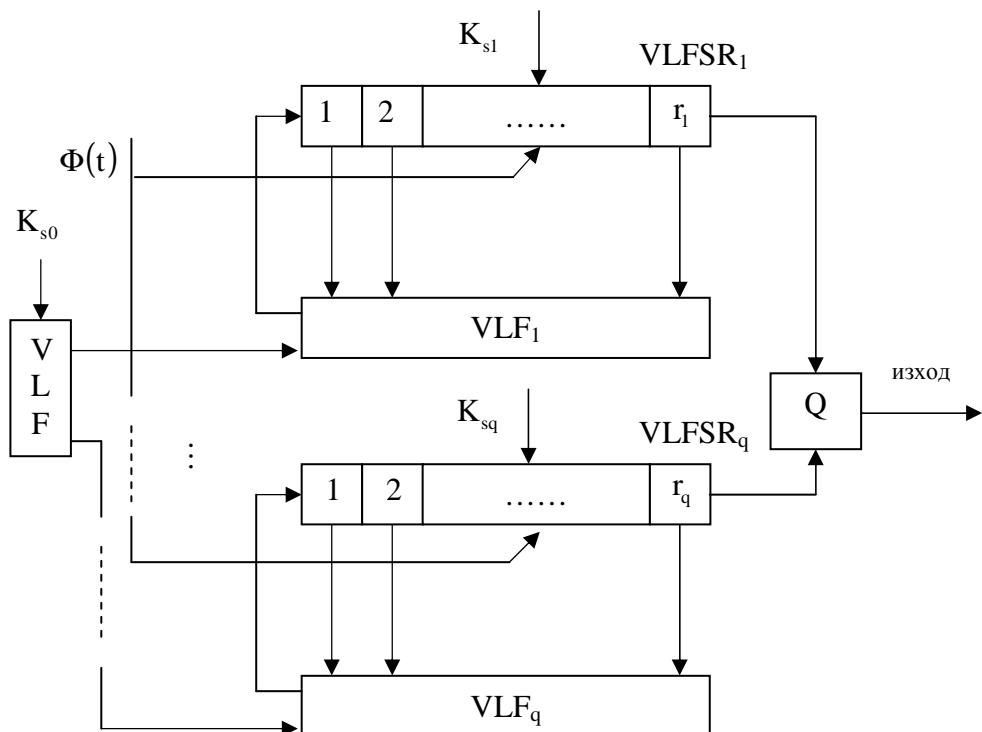
От своя страна, преместващите регистри с линейни обратни връзки (LFSR) могат да се разглеждат като частен случай на линейните последователностни машини - **ЛПМ** (**LSC** - Linear Sequential Circuits), които са достатъчно добре описани в литературата [3,5 и др.].

Всяка ЛПМ се характеризира с входен вектор, изходен вектор, вектор на състоянията и характеристични матрици [3,5 и др.]. Въздействията, приложени към входа, приемат стойности от крайното поле на Галуа **GF(p)**, а самите ЛПМ се изграждат от три различни типа елементарни съставляващи, всяка от които изпълнява една от следните функции [3,5 и др.]: събиране на входните сигнали, умножение на входния сигнал с константа или забавяне на входния сигнал на един такт [операциите събиране и умножение се реализират в съответствие с правилата в полето на Галуа **GF(p)**]. В случай, че изходите и състоянията на

ЛПМ не зависят от входните въдействия, такива ЛПМ се наричат автономни ЛПМ (АЛПМ). При двоично поле на Галуа ($p=2$) тези АЛПМ се превръщат в отбележаните по-горе **LFSR**, които се използват и за генериране на ключови потоци за поточно шифриране. В това направление съществуват множество разработки, но независимо от това, задачата за създаване на качествени нови генератори на ключови потоци на основата на LFSR продължава да бъде актуална. В тази връзка, в настоящата работа се представят нови обобщени схеми на такива генератори, които могат да бъдат използвани за създаване на множество конкретни реализации. Схемите са базирани на известния генератор Hughes XPD/KPD [виж 6 и др.] и се явяват негови модификации, обобщения и развитие.

Схеми на ключовите генератори

Първата представена схема на генератор на ключови потоци е показана на фиг. 2. Тя се явява обобщение на генератора Hughes XPD/KPD, в който, както е известно, се използва един 61-битов LFSR, 2^{10} полинома за обратна връзка (ПОВ) и 8 нелинейни филтъра за формиране на ключовия поток на изхода. С помощта на секретният ключ се установява началното състояние на 61-битовия регистър и се избира един от възможните ПОВ.

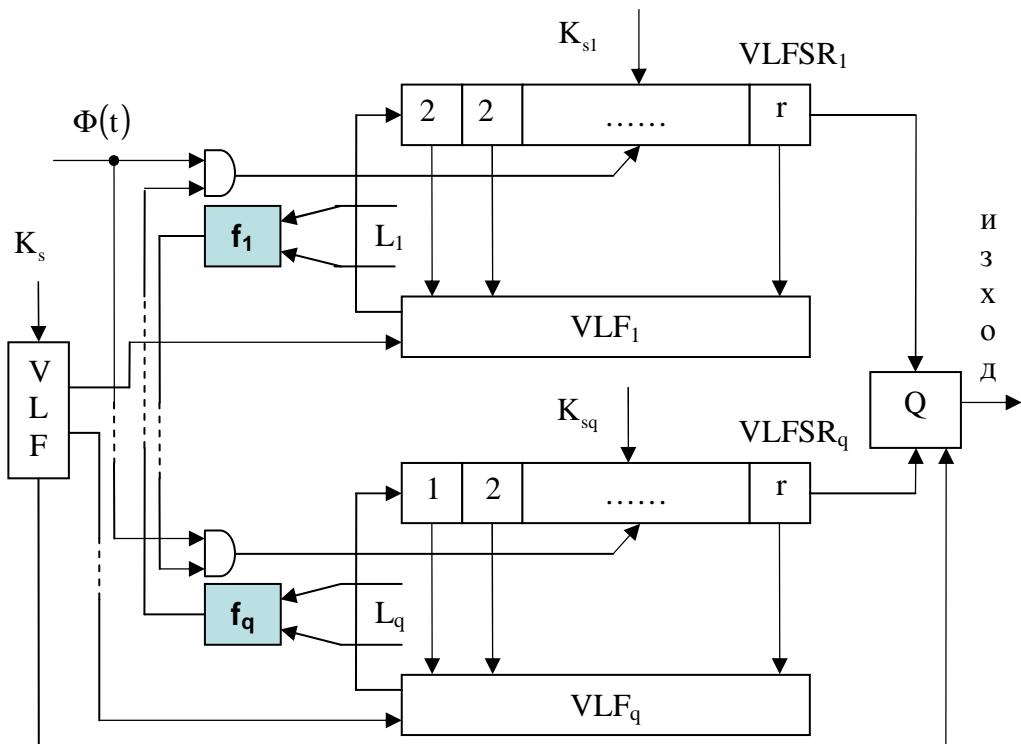


Фиг. 2

В схемата на фиг.2, $\Phi(t)$ - тактираща последователност от импулси, управляваща преместването в регистрите; $VLFSR_1 \div VLFSR_q$ - преместващи регистри с променливи (непостоянни) линейни обратни връзки (Variable LFSR) - броят на тези регистри е q и всеки от тях може да е с различна разрядност; $VLF_1 \div VLF_q$ - множества от полиноми за обратна връзка (ПОВ) за съответните преместващи регистри; VLF (Variable Linear Feedback) - блок за избор на ПОВ (с помощта на този блок от паметите на преместващите регистри се избират конкретни полиноми за всеки отделен регистър в рамките на даден сеанс на връзка).

В горната схема е въведено понятието преместващ регистър с променлива линейна обратна връзка (VLFSR) за обозначаване на тези LFSR, които разполагат с памет за няколко ПОВ. За разлика от тях, LFSR с единствен ПОВ могат да бъдат наречени преместващи регистри с постоянна линейна обратна връзка - CLFSR (Constant LFSR).

Другите обозначения на фиг. 2 са следните: K_{s0} - първата част от секретния ключ, управляваща блока VLF, а $K_{s1} \div K_{sq}$ - останалите части на този ключ, задаващи началните състояния на преместващите регистри; Q - линейна или нелинейна функция, обединяваща изходите на отделните регистри за получаване на окончателната ключова последователност на изхода на генератора.



Фиг. 3

Вижда се, че в схемата на фиг. 2 всеки пореден тактов импулс води до едновременно преместване на един разряд във всички регистри. В следващата схема на ключов генератор, представена на фиг. 3, преместванията в отделните тактове зависят и от съдържанието на определени разряди на регистрите LSFR, задавани от съответните множества ($L_1 \div L_q$). В тези множества могат да се включват само определени или всички разряди, които се използват като входове на линейните или нелинейни филтри ($f_1 \div f_q$). От своя страна, изходите на филтрите участват в управлението на преместванията в различни регистри LFSR. Например, ако изходът на филтър f_q на фиг. 3 в текущия такт е 0, то в следващия такт в регистър VLFSR₁ няма да се осъществи преместване и т.н.

Освен това, в схемата на фиг. 3 е предвидена възможност за избор и между различни възможни функции Q , в зависимост от първата част на секретния ключ K_{s0} .

Тази схема позволява да се направят и допълнителни модификации, като например, да се включи q -разряден преместващ регистър с линейна обратна връзка, съдържанието на разрядите на който да участва, заедно с $\Phi(t)$, в определянето на активните и пасивни тактове на преместване на основните регистри VLFSR₁÷VLFSR_q. Началното състояние на този допълнителен регистър е логично да се задава също от съответна част на секретния ключ.

При положение, че регистрите на фиг. 2 и фиг. 3 са с еднаква разрядност, то полиномите за обратна връзка (ПОВ) могат да бъдат съхранявани, не в разделни памети за отделните регистри, а в обща памет, от която да се избират конкретни полиноми за всеки от сеансите на връзка.

Заключение

Както се вижда, представените по-горе схеми на генератори на ключови потоци за поточно шифриране са обобщени и на тяхна основа могат да се синтезират множество конкретни реализации, всяка от които трябва да се анализира с помощта на известните за целта тестове [4,6 и др.].

Вижда се, че обобщените схеми на фиг. 2 и фиг. 3 могат да бъдат с по-висока линейна сложност от класическия генератор Hughes XPD/KPD, което предполага и по-висока сигурност. Например, ако при атака на грубата сила на Hughes XPD/KPD са необходими $(2^{61}-1).2^{10}$ опита, то за генераторите на фиг. 2 и фиг. 3 този брой опити може да бъде значително по-голям.

Литература

- [1]. Антонов, П., С. Малчев. Криптография в компютърните комуникации. В., 2000, 315 с.
- [2]. Авдошин С. М., А. А. Савельева. Криптоанализ: современное состояние и перспективы развития. //Информационные технологии, № 3, 2007. [http://se.hse.ru/data/712/315/1234/Авдошин.Савельева_Криптоанализ.pdf].
- [3]. Гилл, А. Линейные последовательностные машины. Анализ, синтез и применение. Пер. с англ. Москва, Издательство “Наука”, 1974, 287 с.
- [4]. Игоничкина, Е. В. Статистический анализ поточных шифров. //Доклады ТУСУРа, 2 (18), часть 1, июнь 2008, с. 56-57.
- [5]. Фараджев, Р. Г. Линейные последовательностные машины. Москва, Издательство “Советское радио”, 1975, 248 с.
- [6]. Schneier, B. Applied Cryptography. Protocols, Algorithms and Source Code in C. John Wiley & Sons, Inc., 1996, 758 p.

За контакти:

Доц. д-р инж. Петър Цветанов Антонов
Катедра “Компютърни науки и технологии”
Технически университет - Варна
E-mail:peter.antonov@ieee.org
Рецензент: проф. дтн. инж. Г. Маклаков



ДИЗАЙН НА WEB БАЗИРАНА УЧЕБНА СРЕДА СПОСОБСТВАЩ УСВОЯВАНЕТО НА ЗНАНИЯТА

Бойка Ж. Градинарова

Резюме: При дизайна на учебния материал предназначен за обучение подпомогнато от компютри все по-голямо влияние оказват съвременните информационни технологии и Web базираните системи. Съвременните образователните системи променят своето ударение от тренинга към ученето. Курсовете които се разпространяват също се променят от курсове с голямо и негъвкаво съдържание към курсове където съдържанието се представя чрез последователни малки учебни обекти. Концепцията на малките учебни обекти и съвременните Web технологии подпомагат многократната употреба, оперативната съвместимост и достъпността до критериите за дизайн използвани понастоящем от повечето учебни системи. В статията се доразвиват критериите за дизайн включващи концепцията, подпомагаща адаптирането на съдържанието към нуждите на обучаемите. Представя се модел на учебната среда съдържащ откриването на знания като инструмент за автоматично изучаване на модели на учебно поведение от профилите и историята на студентите.

Instructional Design of Web-Based Learning Environments

Boyka Gradinarova

Abstract—Now days instructional design for computer-assisted learning has been strongly influenced by advances in information technology, Internet and Web-based systems. The emphasis of educational systems has shifted from training to learning. The course delivered has also been changed from large inflexible content to sequential small learning objects. The concepts of learning objects together with the advanced technologies of Web and communications support the, interoperability, and accessibility design criteria currently exploited by most learning systems. These concepts enable just-in-time learning. We propose to extend theses design criteria further to include the ability of learning concept that will help adapting content to the needs of learners. The ability of learning concept offers a better personalization leading to the creation and delivery of course content more appropriate to performance and interest of each learner. In this paper we present a new framework of learning environments containing knowledge discovery as a tool to automatically learn patterns of learning behavior from learners' profiles and history.

1. Introduction

Knowledge mining is the discovery of hidden knowledge stored possibly in various forms and places in large data repositories. Knowledge is a valuable asset to most organizations as a substantial source to enhance organizational competency [1]. Researchers and practitioners in the area of knowledge management view knowledge in a broad sense as a state of mind, an object, a process, an access to information, or a capability [2]. People in knowledge management community use the term knowledge assets to refer to any organizational intangible assets related to knowledge such as know-how, expertise, intellectual property. Knowledge assets can be stored in data repositories either in implicit or explicit form [12], [3]. Explicit knowledge can be managed through the existing tools available in the current database technology. Implicit knowledge, on the contrary, is harder to achieve and retrieve. Specific tools and suitable environments are needed to extract such knowledge.

Recently, Li and Chang [6] proposed a model to systematically mine and manage useful knowledge in R&D organizations. This model is a user-centric approach to manage knowledge objects in a cyclic manner along the transitional processes, i.e. knowledge sharing, knowledge processing, knowledge presentation, knowledge capture and knowledge discovery. The authors [6] designed this model to application. We, nevertheless, consider the model sufficiently captures generic idea in knowledge mining and management to be useful in other context as well. We propose to apply the model to manage knowledge in the web-based learning environments.

A top-level of knowledge management model in web-based learning environments is shown in Fig. 1. From the data and knowledge repositories, the process of knowledge discovery has been applied to acquire knowledge objects that will be subsequently processed in the indexing and mapping stage. This stage supports the search for suitable contents presented to learners. Performance and learners' preference are then captured to store as history and learnable objects to be later used in the discovery stage.

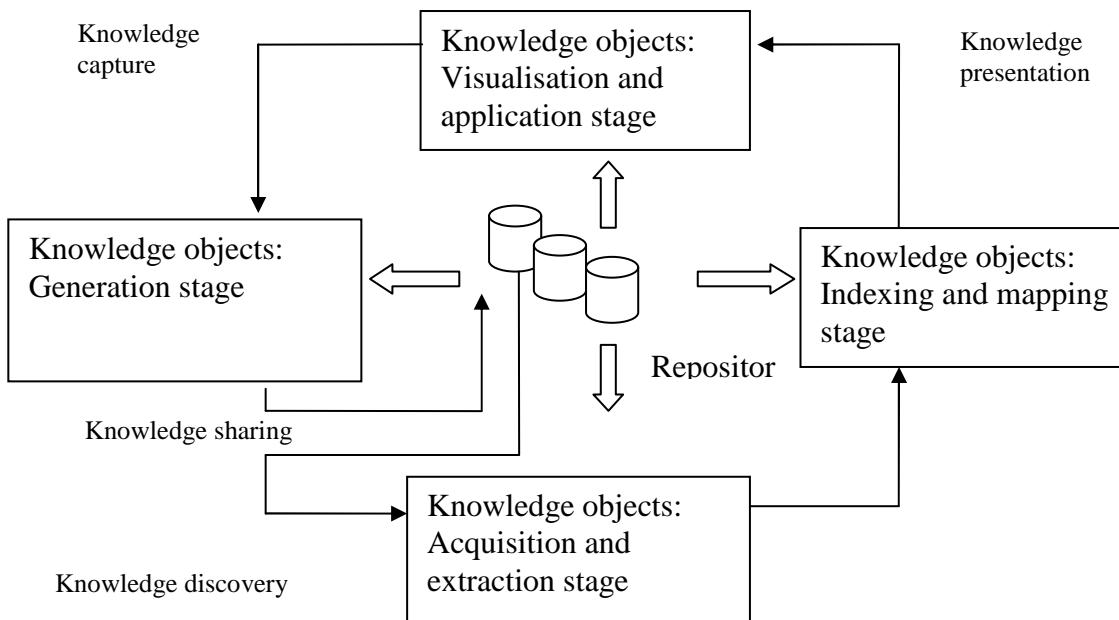


Fig. 1. A top-level of knowledge management model in web-based learning

The rest of this paper is organized as follows. Section 2 presents the concept of web-based learning environments. Section 3 discusses in detail our framework of integrating knowledge mining capability to the web-based learning environments to enhance content management ability. Section 4 concludes the paper.

2. WEB-Based Learning Environments

Web-based learning environments refer to computer-based and computer supported education and training systems exploiting Web as the representation and delivery medium. A web-based learning system is a state-of-the-art in the field of e-learning due to the advantages of reusability, interoperability and accessibility provided by the Internet technology [2]. Reusability,

interoperability and accessibility are major concepts in the development of modern learning environments studied and extensively proposed by many researchers [7].

Davis et al. [3] proposed learning environments to work with didactic materials. Their design is based on interoperability and reusability concepts. Valderrama et al. [16] designed web-based education systems as a multi-agent architecture working with intelligent reusable learning objects. Saddik et al. [8] conducted a project to develop a web-based adaptive hypermedia learning system called Multibook. Fong et al. [4] also studied multimedia courseware in a wireless campus environment. Gunther et al. [6] proposed the concept of sharing statistical computing modules in a web-based system.

Reusability concept offers facility for course content developers to avoid rewriting on updated materials and duplicating effort from the design process. With the interoperability and accessibility concepts, learners are not restricted to a specific content system in a single and static site. The ability to learn contents from multiple sources on different kinds of machines ranging from notebook and PC to handheld and PDA devices requires standardizing the learning contents.

SCORM (Shareable Content Object Reference Model) [5] is currently adopted by a lot of industrial and educational organizations as a set of standards to specify course structure and content delivery process. SCORM is a collection of standards that enable interoperability, accessibility and reusability of distributed heterogeneous web-based learning systems. Reusability in SCORM model can be achieved through the concept of learning objects.

A learning object is defined as an independent collection of content and media elements. A learning object is thus the granule of content. Its size can be varied from a word, a sentence, a paragraph, to a whole chapter of content material. The general guideline on designing learning content is that each learning object should cover an instructional learning objective. SCORM describes the way content objects should be created, discovered and aggregated into more complex learning objects and then organized into a sequence of delivery.

Besides SCORM which is the product of Advanced Distributed Learning (ADL) organization, there exist other standards such as Instructional Management Standards (IMS) [8]. Standardization in web-based learning is an evolving aspect and its extension has been proposed in several work [5], [7]. This paper is yet another extension of a learning model. We propose to raise the learning objects as commonly referred to by SCORM standard to the level of knowledge objects. Our proposed learning model emphasizes the inclusion of knowledge discovery module to discover new knowledge that can be exploited to better planning and individualizing course content.

Individualizing or personalization concept has emerged from the observation that a one-fits-all style of instructional design and teaching strategy is unlikely to lead to learning effectiveness. This is due to the fact that each learner is different from others in terms of background, learning goal, learning skill and learning ability. Instructional methods that match learning style individually will be the most effective learning environment.

Personalization thus leads to adaptive learning environments in which individual learners can be uniquely identified. A learning system designer can take a step further to make the learning process autonomous by providing the environments that learners can take control over their own organization of learning. In distance learning where course is to be delivered to distributed learners, collaborative learning is necessary. Collaborative learning requires communication technologies either in the form of email, shared workspaces, or video conferencing in order to provide learners

with the means of group-oriented learning and discussion exchange between instructor and group of learners.

To design a successful web-based learning system, the designer has to aware of the four main factors as graphically displayed in Fig. 2. These factors form the basis for most learning environments including ours as will be discussed in Section 3.

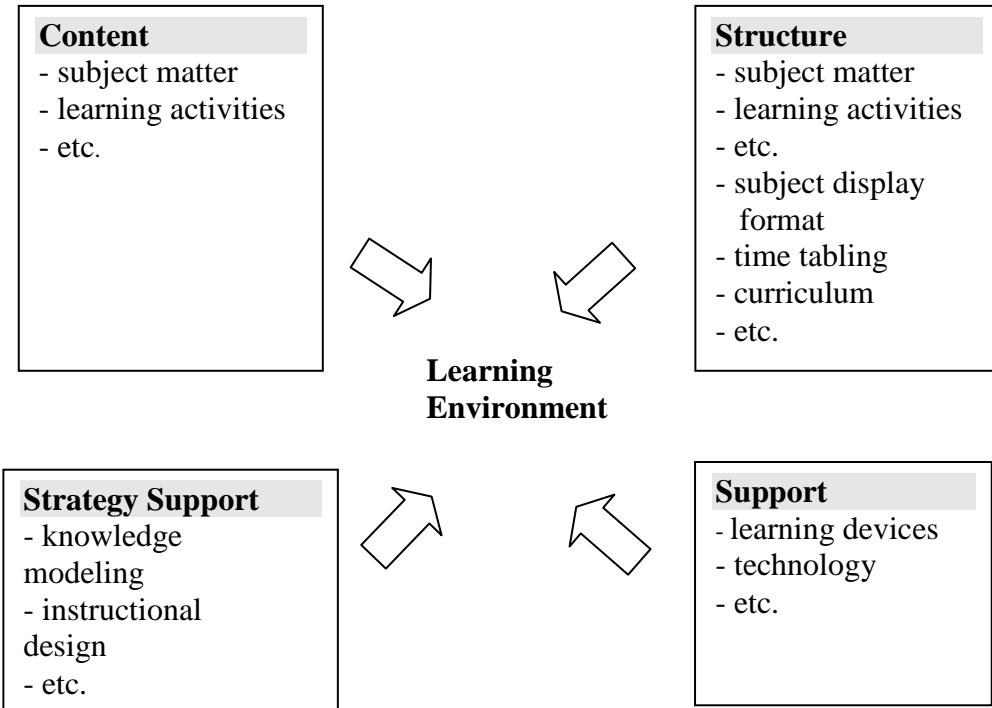


Fig. 2 A group of major factors influencing the success of a learning system

3. Knowledge discovery and web-based learning

The design of our learning model is on the basis of learning content management system [1], [9]. There are three main modules in our proposed framework (as shown in Fig. 3), i.e. learning management, content management, and knowledge management modules. Our framework is proposed to support web-based learning with several scheme of learning. including adaptive, autonomous, and collaborative learning.

Learning management modules: provide the following capabilities:

- support instructors to post syllabi, class schedules,
- assignments, lecture notes, slides and other supplemental materials for learners to access via Web browsing tool.
- support instructors to conduct assessments in various forms such as online tests, surveys, quizzes using a variety of standard question formats, e.g. multiple choice, true/false, essay, short-answer, matching, etc.

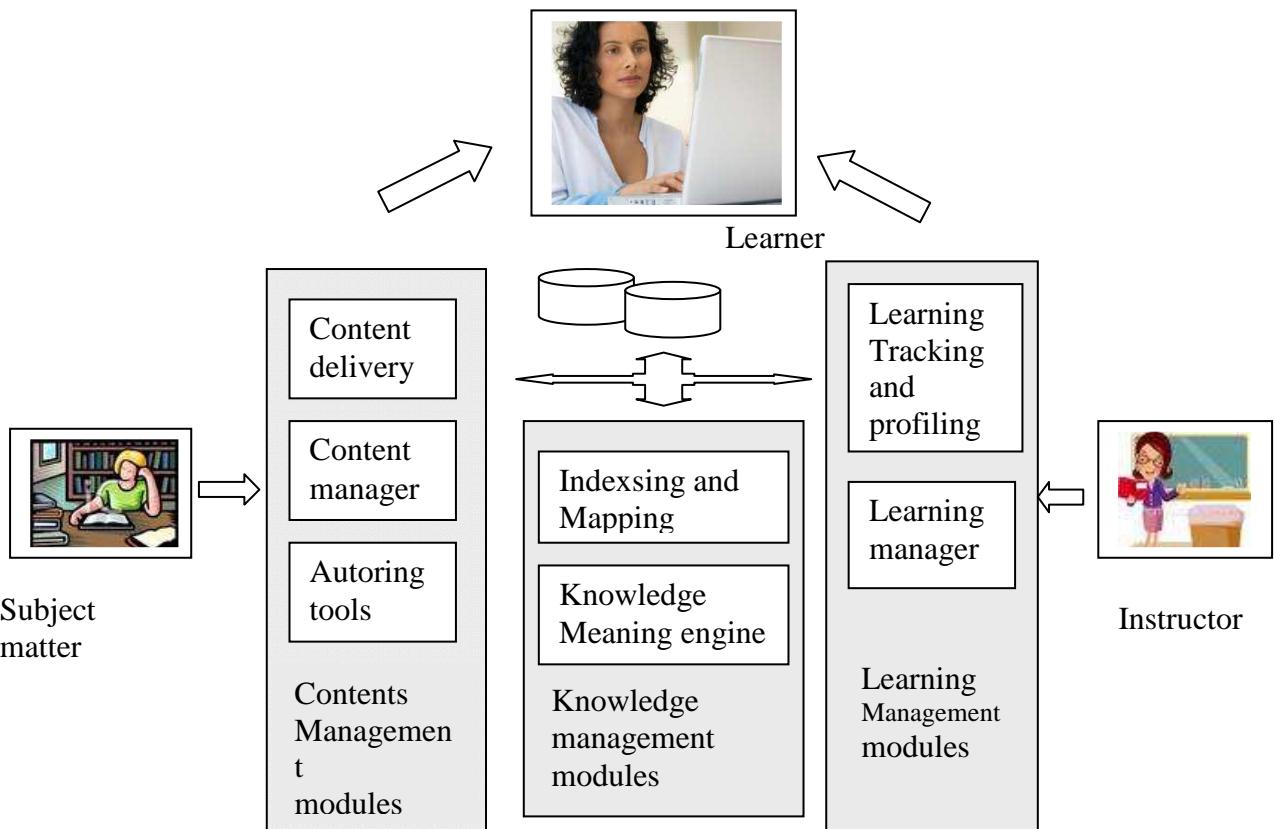


Fig. 3 A framework of a web-based learning environment with main modules for learning management, content management and knowledge management

- support learners to submit assignments remotely either as file upload or interactive through Web interface.
- provide profiling tool to collect personal data of learner and tracking tool to observe learners' actions including like and dislike information.
- provide matchmaking tool to compare the created profile with the available content.
- support instructors and learners in collaborative discussion on assignments and course content.

Content management modules: provide the following capabilities:

- support the content developers in importing and exporting content through the authoring tools.
- support the content manager in individualize the presented content.
- support the content manager in archiving and versioning the content.
- interface with learning management modules in getting desired form for delivered content.
- interface with data repository containing learners' personal information and other metadata including knowledge assets created by knowledge management modules and apply these data in creating personalized sequence of content material suitable for each learner.

Knowledge management modules: provide the following capabilities:

- discover valuable knowledge assets from the data repository containing learners' personal data, tracked data of learners' performance and behavior, and data related to content sequences that were presented in the past with the evaluation results according to that content sequence.
- support the indexing and mapping of knowledge assets that are discovered by the knowledge mining engine.

The major component of learning management modules is learning manager which acts as a conductor controlling and synchronizing every component within the modules. The manager component is also responsible for interfacing with the storage. This is also the case for the content manager in content management modules. The authoring tools in content management modules support creation of all types of digital content materials such as word documents, spreadsheet data, pictures in standard formats, video content, animation and multimedia data.

For the knowledge management modules, knowledge mining engine is responsible for the synchronizing process. Indexing and mapping is a component for storing and searching knowledge assets to be used in the learning process.

4. Conclusion

Recent developments in information and communication technology certainly influence the design and implementation of educational systems. An emergence of the World Wide Web in the early 1990s has resulted in a substantial change in both the content representation and the delivery mechanism. The XML (eXtensible Markup Language) is a currently acclaimed technology as the main content representation format. The main advantage of XML over other hypertext languages such as HTML is its property of being interoperable data interchange format. This language technology helps improving the courseware design concept toward interoperability, which is one of the major design criteria for most software products these days.

Improvements in hardware and communication technology such as mobile microprocessors and wireless communication have evolved learning environments to the mobile and distributed platforms. Such information infrastructures provide learners a remote access to experts and distributed resources. Web browser is a software device offering learners with open and flexible accessibility to distant course contents. The organization of learning resources has also been changed from creating and delivering large inflexible course content to producing database-driven learning objects that can be reused, searched and modified independent of the delivery media.

Interoperability, accessibility and reusability are therefore the main design concepts of current instructional systems. These concepts have shifted the learning paradigm from static learning to collaborative, adaptive and autonomous learning. Collaborative learning allows direct contact between instructor and group of learners through the communication technologies ranging from email and share workspaces to video conference systems. Adaptive and autonomous learning allows learners to take control over their own organization of learning pace and scheme.

In this paper we argue that with the matured technology of knowledge discovery in databases or KDD, the integration of knowledge discovery capability to the creation, delivery and management of learning objects should be the next step in e-learning. Our proposed architecture of learning environments will enable the convergence of e-learning with knowledge management. A

repository containing learner-related materials is a valuable source of knowledge to support personalization information for independent learners.

References

- [1] Chapman B. and B. Hall, Learning Content Management System. Brandonhall.com, New York, 2005.
- [2] Davis L., R.F. Gamble, and S. Kimsen, “A patterned approach for linking knowledge-based systems to external resources,” IEEE Transactions on Systems, Man and Cybernetics, Part B, 34(1), pp. 222-233, 2004.
- [3] George S. and H. Labas, “E-learning standards as a basis for contextual forums design,” Computers in Human Behavior, 24, pp. 138-152, 2008.
- [4] Huang Y.M., J.N. Chen, T.C. Huang, Y.L. Jeng, and Y.H. Kuo, “Standardized course generation process using dynamic fuzzy petri nets,” Expert Systems with Applications, 34, pp. 72-86, 2008.
- [5] IMS, “IMS learning design information model version 1.0,” IMS Global Learning Consortium Inc., Retrieved March 2008 from <http://www.imsglobal.org/learningdesign/index.html>.
- [6] Li S.T. and W.C. Chang, “Exploiting and transferring presentational knowledge assets in R&D organizations,” Expert Systems with Applications, 2007, doi:10.1016/j.eswa.2007.10. 024.
- [7] Ruppel C.P. and S.J. Harrington, “Sharing knowledge through intranets: A study of organizational culture and intranet implementation,” IEEE Transactions on Professional Communication, 44(1), pp. 37-51, 2001.
- [8] Saddik A.E., S. Fischer, and R. Steinmetz, “Reusable multimedia content in Web based learning systems,” IEEE Multimedia, 8(3), pp. 30-38, 2001.
- [9] Valderrama R.P., L.B. Ocana, and L.B. Sheremetov, “Development of intelligent reusable learning objects for web-based education systems,” Expert Systems with Applications, 28(2), pp. 273-283, 2005.

За контакти:

гл.ас. д-р инж. Бойка Жекова Градинарова
катедра “Компютърни науки и технологии”

Технически университет - Варна
E-mail: Bgradinarova@hotmail.com

Рецензент: доц. д-р инж. Н. Рачева, ТУ – Варна

СПЕЦИФИКАЦИЯ IMS QTI И АНАЛИЗ ТЕСТОВЫХ ЗАДАНИЙ С КОРОТКИМ ОТВЕТОМ НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ РЕГУЛЯРНЫХ ВЫРАЖЕНИЙ

Валерий И. Давыдов, Петр Н. Фомин

Резюме: В статье рассматривается классификация, достоинства и недостатки типов тестовых заданий, вопросы проверки заданий с коротким ответом на основе использования регулярных выражений.

Specification of IMS QTI and Analysis Test Tasks With a Short Answer on the Basis of the Use of Regular Expressions

Valery I. Davyдов, Peter N. Fomin

Abstract: The article is devoted to the classifications, advantages and drawbacks of test tasks types, verification questions of test tasks with a short answer on the basis of use of regular expressions.

1. Введение

Для оценки результатов самостоятельной работы студентов широко используются тестовые задания различных типов.

К тестам предъявляются следующие основные требования:

- четко определенные цели;
- валидность, оперативность, полнота, непрерывность, объективность, надежность, однозначность и дифференцирующая способность;
- разнообразие типов тестовых заданий;
- возможность реализации различных видов контроля (самоконтроль, входной, текущий, модульный и итоговый контроль);
- возможность использования в тестах средств мультимедиа;
- мониторинг итогов работы студентов, коррекция тестовых заданий;
- соответствие тестов международным спецификациям и стандартам.

Рассмотрим классификацию типов тестовых заданий. Особое внимание уделим анализу заданий с коротким ответом на основе использования регулярных грамматик.

2. Классификация тестовых заданий

Международный консорциум IMSGlobal (IMS Global Learning Consortium) разработал спецификацию IMS Q&TI, которая наиболее полно отвечает указанным требованиям. В спецификации определены более 20 типов тестовых заданий, классификация которых приведена в табл. 1.

Типы тестовых заданий классифицируются:

- по типу ответа:
 - ввод логического идентификатора;
 - ввод координат X-Y;
 - ввод текста, ввод числа;

- установка соответствия логических групп.
- по формату представления ответа:
 - одиночный;
 - множественный;
 - упорядоченный.

Таблица 1 Классификация типов тестовых заданий (спецификация IMS QTI)

Тип ответа	Формат представления		
	Одиночный (Single)	Множественный (Multiple)	Упорядоченный (Ordered)
Логический идентификатор (Logical Identifier – LID)	Единственный выбор (Multiple choice) Истина/ложь (True/false) Слайдер (Slider)	Множественный выбор (Multiple response)	Упорядочить объекты (Order objects) Соединить точки (Connect points) Установить соответствие объекта (Match object) Упорядочить объект (Order object) Переместить объект (Drag object) Переместить “мишень” (Drag target)
Координаты X-Y (Coordinates)	Указать область (Image hot spot)	Упорядочить объекты (Order objects)	Соединить точки (Connect points)
Ввести текст (String – STR)	Заполнить бланк (Fill-in-blank) Выбрать текст (Select text) Короткий ответ (Short answer) Эссе (Essay)		
Число (Numerical–NUM)	Заполнить бланк (Fill-in-blank) Слайдер (Slider)		
Логические группы (Logical Groups –GRP)	Слайдер (Slider) Логические группы (Logical Groups –GRP) Установить соответствие объектов (Match objects) Переместить объект (Drag object) Переместить “мишень” (Drag target)	Установить соответствие объектов (Match objects) Переместить “мишени” (Drag targets) Переместить объекты (Drag objects)	Установить соответствие объектов (Match objects) Упорядочить объекты (Order objects)

В отечественной литературе различают тестовые задания:

- закрытой формы;

- на установление соответствия;
- на установление правильной последовательности;
- открытой формы.

Рассмотрим более подробно некоторые типы заданий, которые широко используются в системах дистанционного обучения.

3. Тестовые задания закрытого типа

Закрытая форма тестовых заданий наиболее известна и чаще всего употребляется в практике тестирования. В таких заданиях приводится несколько ответов, из которых хотя бы один правильный. Многие разработчики тестов ошибочно понимают под тестом исключительно эту форму заданий. К достоинству закрытой формы следует отнести простоту технической реализации автоматической проверки заданий. Существенным недостатком закрытой формы является высокая вероятность угадывания ответа. Если считать ответы равновероятными, то в “задании с ответом Истина/Ложь (True/false)” вероятность угадывания правильного ответа равна $p = 0,5$. Учитывая столь высокую вероятность угадывания ответа, такие задания не рекомендуется использовать в педагогических тестах.

При использовании тестового задания “выбор одного ответа из N ” (единственный выбор – Multiple choice в спецификации IMS QTI) вероятность угадывания ответа $p = 1/N$. Для снижения вероятности угадывания ответа, такие задания рекомендуется использовать при $N > 5$.

При использовании тестового задания “множественный выбор k правильных ответов из N ” (множественный выбор – Multiple response в спецификации IMS QTI), если студент указывает не все правильные ответы, существует два подхода к оценке результата:

1. Частично правильный.
2. Неправильный.

<p>Какие из записей являются 16-ричными константами Visual Basic?</p>	<p>Ответ: <input type="checkbox"/> a. &HB56FA3 <input type="checkbox"/> b. &H56FA34 <input type="checkbox"/> c. &h45ta2K3 <input type="checkbox"/> d. &H37FA34 <input type="checkbox"/> e. &H4fA1eh <input type="checkbox"/> f. &H56FA3</p>
Ответить	

Рис. 1. Пример тестового задания типа «множественный выбор»

Наиболее часто используется второй подход. В этом случае вероятность случайного угадывания правильных ответов, если k неизвестно студенту, определяется выражением $p = 1/ 2^N$. При $N = 4$ вероятность угадывания ответа $p = 0,063$, что можно считать допустимым при тестировании.

Рассмотрим пример. Предположим, что мы хотим проверить знание студентом правил построения лексем языка Visual Basic (например, 16-ричной константы). При использовании теста типа “множественный выбор” интерфейс задания может иметь вид, представленный на рис. 1 (правильные ответы выделены жирным шрифтом).

Вероятность случайного угадывания ответа в тестовых заданиях типа “установление соответствия двух множеств” (Match objects в спецификации IMS QTI), состоящих из N объектов, равна $p = 1/ N!$. При $N = 4$ вероятность угадывания $p = 0,042$, что меньше чем в предыдущем типе.

Вероятность случайного угадывания ответа в тестовых заданиях “установление правильной последовательности k объектов из N ” (Order objects в спецификации IMS QTI) существенно меньше предыдущего варианта и определяется следующим выражением:

$$p = 1 / \sum_{k=2}^N N! / (N - k)! \quad (1)$$

Рассмотренные выше типы тестовых заданий достаточно просты в технической реализации, однако, для эффективной реализации педагогических задач их явно недостаточно, тк они позволяют только приближенно оценить степень усвоения изучаемого материала. Эти типы заданий, как правило, позволяют реализовать тесты 1 и 2 уровня усвоения [2] и не позволяют проверить ход и логику мышления обучаемых.

4. Тестовые задания открытого типа

Тестовые задания открытой формы (в спецификации IMS QTI: эссе (Essay), заполнить бланк (Fill in-blank), короткий ответ (Short answer)) обладают следующими достоинствами:

- достаточно высокой надежностью, т.к. случайное угадывание ответа невозможно;
- возможностью реализации тестов 3 уровня.

Задания открытой формы имеют следующие недостатки:

- сложность технической реализации автоматической проверки;
- сложность составления заданий в случае многовариантности правильного ответа.

Эти типы текстовых заданий некоторые авторы относят к заданиям со свободно-конструируемым ответом. С этим трудно согласиться, так как только эссе полностью соответствует этому типу. Техническая реализация автоматической проверки ответа в общем случае для любой предметной области очень сложна, т.к. требует выполнения проверки ответа на естественном языке с использованием семантического, синтаксического и лексического анализаторов. По этой причине проверка заданий типа эссе в большинстве систем дистанционного обучения не выполняется автоматически. Такая проверка выполняется преподавателем “вручную”, что существенно усложняет и замедляет процесс тестирования.

При использовании тестового задания типа Fill in-blank обучаемый должен ввести в свободное поле:

- строку текста (чаще всего одно слово, реже фразу);
- результат вычисления заданного выражения или фрагмента программы.

Проверка ответа основана на сравнении с эталонами правильных ответов.

При использовании ввода строки в свободное поле можно построить тест, который будет требовать от обучаемого выполнения глубокого анализа задания, принятия решения с учетом контекста, окружающего свободное поле. Следует помнить, что эталоны должны содержать все возможные варианты ответов, т.к. в противном случае неизбежны ошибки проверки правильности ответа.

Тестовое задание с коротким ответом (Short answer) требует от обучаемого ввода одного слова или результата вычисления заданного выражения. Проверка этого типа задания реализуется путем сравнения ответа обучаемого с множеством эталонов правильных ответов. Разработчик тестового задания, как и в предыдущем случае, должен учитывать возможность многовариантности ответов.

5. Проверка тестовых заданий открытого типа с использованием регулярных выражений

Определение множества эталонных ответов можно существенно упростить при использовании регулярных выражений [3], которые содержат мощный, гибкий и

эффективный метод обработки текста. Регулярные выражения содержат два основных типа знаков: текстовые символы и метасимволы. Возможности обработки данных с помощью регулярных выражений определяются набором метасимволов. Наиболее известными метасимволами являются знаки «?» и «*», используемые в файловой системе DOS для представления любого символа или группы символов. Регулярные выражения во много раз расширяют эту основную идею, предоставляя большой набор метасимволов. Обширная нотация для поиска соответствий позволяет выполнять быстрый разбор большого текста и поиск определенных групп символов для извлечения, изменения, замены или удаления текстовых подстрок. Существует несколько спецификаций регулярных выражений, наиболее известные из них: PCRE (Perl Compatible Regular Expressions), POSIX (Portable Operating System Interface for UNIX) и регулярные выражения .NET Framework, которые совместимы с регулярными выражениями языка Perl 5.

Классы регулярных выражений .NET Framework являются частью библиотеки базовых классов, их можно использовать с любым языком, работающим с общеязыковой средой выполнения, например ASP.NET и Visual Studio .NET.

Листинг. Проверка тестовых заданий с использованием регулярных выражений

```
Imports System.Text.RegularExpressions
Public Class Form1
    Inherits System.Windows.Forms.Form
    Private Sub Button1_Click(ByVal sender As System.Object, _
        ByVal e As System.EventArgs) Handles Button1.Click
        Dim Inp, ETALON As String
        Inp = TextBox1.Text
        ETALON = "^&[hH][0-9A-Fa-f]{1,10}$"
        Dim m As Match = Regex.Match(Inp, ETALON)
        If (m.Success) Then
            Label2.ForeColor = Color.Black
            Label2.Text = "Правильно!"
        Else
            Label2.ForeColor = Color.Red
            Label2.Text = "Ошибка!" & vbCrLf & "Попробуйте еще раз!"
        End If
    End Sub
End Class
```

Вернемся к рассмотрению предыдущего примера. Построение лексем языков программирования строго формализованы, поэтому эталон описания 16-ричных констант можно представить в виде регулярного выражения:

$$^&[hH][0-9A-Fa-f]\{1,10\}\$ \quad (2)$$

где: ^, \$ соответствуют началу и концу строки ответа;

[] определяют любой единичный символ, из числа заключенных в скобки. Например, [hH] соответствует буквам “h” или “H”;

0-9 соответствует десятичным цифрам;

A-F соответствует буквам верхнего регистра английского алфавита: A, B, C, D, E, F;

a-f соответствует буквам верхнего регистра английского алфавита: a, b, c, d, e, f;

{1,10} указывает, что последний блок может повторяться не менее одного раза и не более 10 раз.

Листинг программы проверки ответа с использованием регулярного выражения приведен выше.

Программа написана на языке Visual Basic .NET. Для сравнения ответа студента с эталоном используется метод Match класса Regex, который находится в пространстве имен System.Text.RegularExpressions. Пример задания приведен на рис. 2.

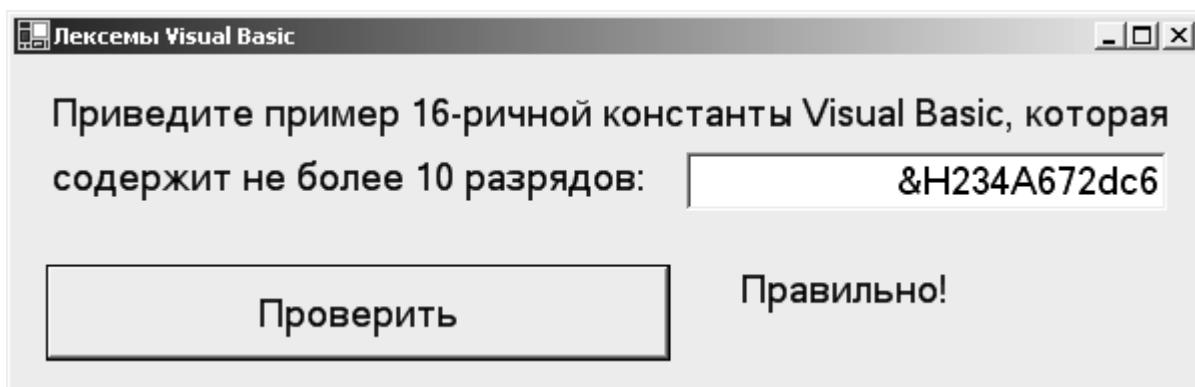


Рис. 2. Пример задания со свободно-конструируемым ответом

При ответе студент должен ввести любое значение 16-ричной константы, которое он выбирает сам, зная правила их построения. Угадывание ответа в этом случае невозможно. После ввода ответа и нажатия клавиши “Проверить”, выполняется процедура обработки события Button1_Click. В случае правильного ответа выводится сообщение “Правильно!” (рис. 2), в противном случае – сообщение “Ошибка! Попробуйте еще раз!”.

6. Заключение

Использование регулярных выражений возможно и в других тестах открытого типа, что позволяет реализовать задания 3 уровня усвоения, в которых четко определена цель, но не совсем ясно каким образом она может быть достигнута. Студент должен выбрать способ решения задания, применив ранее усвоенные навыки, умения и знания, сформулировав ответ в свободно-конструируемом виде. Такие задания требуют от студентов продуктивных действий, самостоятельного обнаружения дополнительных (скрытых) сведений, необходимых для их решения. При этом студенты усваивают субъективно новую информацию в ходе самостоятельного выполнения тестовых заданий.

Рассмотренный подход может быть достаточно легко реализован в системах дистанционного тестирования с использованием платформ ASP.NET и PHP, которые широко используются для создания Web-приложений.

Литература

- [1] IMS Q&TI: ASI Information Model Specification Final Specification. Version 1.2
http://www.imsglobal.org/question/qtiv1p2/imsqti_asi_infov1p2.html
- [2] Беспалько В.П. Основы теории педагогических систем: проблемы и методы психолого-педагогического обеспечения технических обучающих систем. Воронеж, ВГУ, 1977. 304 с.
- [3] Фридл Дж. Регулярные выражения. 2 издание. (Серия “Библиотека программиста”). СПб, Изд-во “Питер”, 2003, 464 с.

Для контактов:

доц. Валерий Давыдов

Кафедра системного программного обеспечения

Национальный политехнический университет – Одесса

E-mail: dvali@ukr.net

Рецензент: доц. д-р инж. Н. Николова, ТУ – Варна

УРОВНЕВАЯ МОДЕЛЬ ВЗАИМОДЕЙСТВИЙ В ЭЛЕКТРОННОМ ОБУЧЕНИИ

Вера В. Любченко

Резюме: В работе предложена модель взаимодействий обучаемого и обучающей среды в электронном обучении. Эта 4-уровневая модель основана на принципах эталонной модели OSI и стратовой модели взаимодействий. Рассмотрены свойства предложенной модели взаимодействий.

Layer Model of e-Learning Interactions

Vera V. Liubchenko

Abstract: The paper offers one model of interactions between learner and learning environment of e-learning. This 4-layer model is based on the OSI reference model and stratified interaction model principles. The properties of proposed interactions model are examined.

1. Введение

Процесс обучения осуществляется между двумя системами: обучающей средой и системой обучаемых. Между этими системами происходит либо одно, либо двунаправленный обмен информацией, в результате которого происходит целенаправленное овладение обучаемым знаниями, умениями и навыками. Задачей обучающей среды — системы управления контентом, системы управления обучением, электронных учебников и т.п. — является создание адекватных условий обучения.

Для создания эффективных обучающих сред необходимо учитывать характеристики механизма обмена информацией в процессе обучения. Для этого целесообразно разработать единую модель взаимодействий, которая будет инвариантна к форме обучения и технологии реализации обучающей среды. Эта модель должна описать правила и процедуры обмена информацией между обучающей средой и системой обучаемых в процессе обучения.

Основой для разработки модели взаимодействий в электронном обучении может быть эталонная модель OSI [1] и стратовая модель взаимодействия пользователя и поисковой системы [2].

2. Эталонная модель OSI

Базовая эталонная модель взаимодействия открытых систем (OSI — Open Systems Interconnection Basic Reference Model) является абстрактной сетевой моделью для коммуникаций и разработки сетевых протоколов [1].

Основополагающим объектом модели OSI является уровень, включающий в себя совокупность функций для решения ряда задач (рис. 1). Каждый из уровней взаимодействует с выше- и нижестоящим уровнем одной системы в соответствии с принципом автономности, который заключает в том, что изменение одного уровня не должно приводить к изменению других уровней. Выполнение данного принципа реализуется через единственную точку взаимодействия уровня.



Рис.1. Модель уровня

Взаимодействие между системами осуществляется посредством протокола и интерфейса. Протоколом называют правила, описывающие последовательность и формат сообщений одного уровня, передаваемых из одной системы в другую. Интерфейсом называют правила передачи сообщений через точку взаимодействия в рамках одной системы.

Уровневая модель исключает прямую связь между соответствующими уровнями разных систем, равно как и связь не смежных уровней одной системы. Поэтому, каждый уровень системы А использует услуги, предоставляемые ему нижестоящим уровнем, чтобы осуществить связь с соответствующим ему уровнем системы В.

Каждый уровень имеет заранее заданный набор функций, которые он должен выполнить для проведения связи.

Прикладной уровень обеспечивает услугами прикладные процессы, лежащие за пределами масштаба модели OSI, идентифицирует и устанавливает наличие предполагаемых партнеров для связи, синхронизирует совместно работающие прикладные процессы, а также устанавливает и согласовывает процедуры устранения ошибок и управления целостностью информации.

Представительный уровень отвечает за то, чтобы информация, посланная из прикладного уровня одной системы, была читаемой для прикладного уровня другой системы, при необходимости осуществляя трансляцию между множеством форматов представления информации.

Сеансовый уровень устанавливает, управляет и завершает сеансы взаимодействия между прикладными задачами, состоящие из диалога между двумя или более объектами представления.

Транспортный уровень предоставляет услуги по транспортировке данных, обеспечивает механизмы для установки, поддержания и упорядоченного завершения действия каналов, систем обнаружения и устранения неисправностей транспортировки и управления информационным потоком.

Сетевой уровень обеспечивает возможность соединения и выбор маршрута между двумя конечными системами.

Канальный уровень обеспечивает надежный транзит данных через физический канал, решает вопросы физической адресации, топологии сети, линейной дисциплины, уведомления об ошибках, упорядоченной доставки блоков данных и управления потоком информации.

Физический уровень определяет электротехнические, механические, процедурные и функциональные характеристики установления, поддержания и разъединения физического канала между конечными системами.

Несмотря на то, что эталонная модель предназначена для описания взаимодействия абстрактных открытых систем, она довольно жестко привязана к описанию сетевого взаимодействия. Выбор числа уровней и определение задач, которые они должны выполнять, во многом определяется необходимостью решения проблемы связанной с обеспечением совместимости оборудования. Для описания взаимодействия обучающей среды и обучаемого

ряд задач не являются характерными, поэтому в исходной формулировке модель OSI использовать не целесообразно.

3. Стратовая модель взаимодействия пользователя и поисковой системы

Стратовая модель взаимодействия пользователя и поисковой системы основана на том, что с точки зрения процесса взаимодействия когнитивных подсистем поиск может быть представлен как серия коммуникативных актов между стратами (уровнями), направленных на решение основной задачи и взаимную адаптацию партнеров [2]. Страты партнера-человека в процессе взаимодействия претерпевают следующие изменения (рис. 2):

- в поверхностной страте определяется характер процесса общения (в случае задачи поиска выполняется применение различных тактик поиска, формулирование и переформулирование запроса, добавление и удаление терминов, работа с документами);
- в аффективной страте определяется эмоциональность — удовлетворенность или неудовлетворенность пользователя результатом;
- в когнитивной и ситуативной стратах происходят изменения, связанные с интерпретацией полученной информации, осмысливанием и переработкой, а также изменение локальных задач, возникающих при достижении общей цели.

Наиболее значимой для понимания особенностей восприятия и использования информации является когнитивная страта. Для нее целесообразно разделять три страты взаимодействия с информационной системой:

- поверхностную страту, на которой осуществляется взаимодействие с интерфейсом системы, работа которого обеспечивается аппаратным обеспечением;
- страту процессов, на которой реализуются алгоритмы решения необходимых пользователю задач;
- содержательную страту, на которой обеспечивается доступ к информационным ресурсам.



Рис. 2. Элементы стратовой модели

Данная модель принципиально отличается от эталонной модели OSI своей ориентацией на особенности восприятия информации человеком и возможностью учесть участие человека в информационном процессе. Однако, она не может быть непосредственно использована при моделировании взаимодействий в электронном обучении по двум причинам. Во-первых, модель ориентирована на специфику объекта моделирования — взаимодействия с поисковой системой. Во-вторых, учет средств обеспечения коммуникационных актов выполнен на основе особенностей восприятия и использования информации человеком, что не позволяет говорить о представлении механизма обмена информацией.

3. Уровневая модель взаимодействий обучающей среды и обучаемого

Анализ рассмотренных моделей позволяет сформулировать свойства, которым должна удовлетворять разрабатываемая модель.

1. Основным объектом модели должен быть уровень, включающий совокупность функций для решения ряда четко определенных задач.

2. Выделение уровней модели нужно выполнять на основе модели страт партнера человека в процессе информационного взаимодействия.

Предлагаемая модель взаимодействия состоит из 4-х уровней, каждый из которых может взаимодействовать только со своими соседями и выполнять отведенные только ему функции. Каждый уровень обеспечивает свою часть процесса взаимодействия, зависимость между уровнями минимальна: вышележащие уровни используют результаты работы нижележащих уровней.

Основу модели образует *уровень структуры*, который определяет содержательное наполнение процесса обучения. Исследования показали, что во время обучения на успешность выполнения каждой функции обработки информации влияет структурированность материала и очевидность ассоциативных связей между основными фактами, феноменами, принципами и законами, которые определяют суть информации и учебного материала [3]. Выявление и фиксация структуры учебного материала обеспечивает успех восприятия на всех уровнях.

Уровень сеансов отвечает за поддержку сеансов обучения, определяя частоту и продолжительность взаимодействия между участниками процесса. Например, правила этого уровня определяют рекомендуемую продолжительность работы с электронным учебником каждую неделю.

Анализ структуры позволяет оценить качество различных способов декомпозиции структурных элементов на содержательные модули. Свойства декомпозиции являются одним из главных факторов для определения характеристик сеансов обучения.

Уровень транспорта предоставляет услуги для доставки информации, передаваемой при взаимодействиях участников процесса обучения. Например, правила этого уровня определяют способ взаимодействия (синхронное или асинхронное) в электронном обучении с использованием Интернет-системы управления обучением.

Следует считать устаревшим убеждение, что средства доставки информации определяются формой обучения. В настоящее время все большую популярность приобретает смешанное обучение, использующее различные средства доставки, выбор которых корректнее осуществлять на основе свойств сеансов обучения.

Наконец, самый высокий уровень — *уровень контактов* — непосредственно обеспечивает обмен информацией между участниками процесса, а также отвечает за передачу организационной с точки зрения процесса обучения информации. Например, правила этого уровня определяют технологию проведения промежуточного тестирования в

электронном обучении. Правила этого уровня отвечают также за преобразование учебных материалов в формат, пригодный для использования в процессе обучения. Очевидно, что все правила уровня контактов основываются на результатах нижележащих уровней.

Рассмотрим свойства предлагаемой модели. Основным объектом модели являются уровни, причем работа нижележащих уровней не зависит от работы вышестоящих. При этом работа каждого уровня воздействует на определенные стратегии обучаемого. Можем выделить три уровня воздействия:

- уровень структуры воздействует на когнитивную стратегию, формируя понимание структуры знаний предметной области изучаемого материала, и на аффективную стратегию, обеспечивая эмоции понимания содержания;
- уровни сеансов и транспорта воздействуют на аффективную стратегию, создавая и поддерживая эмоции вовлеченности в процесс обучения;
- уровень контактов воздействует на ситуационную стратегию, обеспечивая понимание обучаемым потребности в получаемых знаниях для решения конкретных прикладных задач. Следует отметить, что такое понимание оказывает влияние также на аффективную стратегию, создавая эмоции озарения (англ. — insight, нем. — Aha-Erlebnis), и на когнитивную стратегию, давая подкрепление сформированному пониманию структуры знаний.

4. Заключение

В работе предложена модель для описания взаимодействий между двумя системами, участвующими в процессе обучения. Эта модель состоит из 4 уровней, каждый из которых может взаимодействовать только со своими соседями и выполнять отведенные только ему функции. Модель инвариантна к форме обучения и технологии реализации обучающей среды.

Предложенная модель может быть использована в процессе выбора технологических решений при создании новой или для оценки эффективности существующей обучающей среды для электронного обучения.

Литература

- [9]. Олифер, В.Г., Н.А.Олифер. Компьютерные сети. Принципы, технологии, протоколы. СПб., Питер, 2007, 960 с.
- [10]. Saracevic T., Spink A., Wu M. Users and Intermediaries in Information Retrieval: What are They Talking About? / Jameson A. Paris C., Tasso C. User Modeling: Proceedings of the Sixth International Conference, UM97. CISIM, 1997.
- [11]. Melezinek A. Ingenieur-pädagogik: Praxis der Vermittlung technischen Wissens. – Wien, New York: Springer, 1999. – 239 s.

Контактный адрес:

доц. к.т.н. Вера Викторовна Любченко

Кафедра системного программного обеспечения

Одесский национальный политехнический университет

E-mail: lvv@edu.opru.ua

Рецензенти: доц. д-р инж. Н. Николова, ТУ – Варна

доц. д-р инж. А. Смрикаров, Русенски университет "Ангел Кънчев"

Изисквания за оформяне на статиите на авторите за списание "Компютърни науки и технологии"

- I. Ръкописите на статията се представят разпечатани в два екземпляра (оригинал и копие) в размер до 6 страници формат А4 и като файл на адрес: Технически университет – Варна, катедра “Компютърни науки и технологии”, ул. Студентска 1, 9010 Варна.
- II. Текстът на статията трябва да включва: УВОД (поставяне на задачата) ИЗЛОЖЕНИЕ (изпълнение на задачата), ЗАКЛЮЧЕНИЕ (получени резултати), БЛАГОДАРНОСТИ към сътрудниците, които не са съавтори на ръкописа (ако има такива), ЛИТЕРАТУРА и адрес за контакти, включващ: научно звание и степен, име, инициали, фамилия, организация, поделение(катедра), e-mail адрес.
- III. Всички математически формули трябва да са написани ясно и четливо (препоръчва се използване на Microsoft Equation). Номерацията на формулите се дава вдясно от началото на реда. Текста на формулите се позиционира в средата на реда. Експоненциалните функции се изписват чрез знака \exp .
 1. Текстът трябва да бъде въведен във файл във формат WinWord 2000/2003 със шрифт Times New Roman. Форматирането трябва да бъде както следва:
 2. Размер на листа - А4, полета - ляво - 20мм, дясно - 20мм, горно - 15мм, долно - 35мм, Header 12.5мм, Footer 12.5mm (1.25cm).
 3. Заглавие - (на български език, размер на шрифта 16, bold).
 4. Един празен ред, размер на шрифта 14, нормален.
 5. Имена на авторите - име, инициали на презиме, фамилия, без звания и научни степени, размер на шрифта 14, нормален.
 6. Два празни реда, размер на шрифта 14.
 7. Резюме -(на български език, размер на шрифта 11) - до 8 реда, нормален.
 8. Заглавие (на английски език, размер на шрифта 12) – bold.
 9. Един празен ред, размер 11.
 10. Имена на авторите (на английски език, размер на шрифта 11) – нормален.
 11. Един празен ред, размер 11.
 12. Резюме -(на английски език, размер на шрифта 11) - до 8 реда, нормален.
 13. Основните раздели на статията (Увод, Изложение, Заключение, Благодарности) се форматират в едноколонен текст както следва:
 - n. наименование на раздел или на подраздел, размер на шрифта 12, удебелен, центриран.
 - o. празен ред, размер на шрифта 12;
 - p. текст, размер на шрифта 12, нормален, отстъп на първи ред на параграф – 10 mm; разстояние от параграф до съседните (Before и After) за целия текст – 0.
 - q. цитиране на литературен източник - номер на източника от списъка в квадратни скоби;
 - r. Номерация на формулите =- дясно подравнена, в кръгли скоби.
 - s. Графики – ако не се побират в една колона от текста е желателно да се разполагат в началото или в края на страницата. Format Layout: In line with text
 - t. литература – всеки литературен източник се представя с: номер в квадратни скоби и точка, списък на авторите (първият автор започва с фамилия, останалите – с име), заглавие, издателство, град, година на издаване.
 - u. За контакти: научно звание и степен, име, презиме (инициали), фамилия, организация, поделение (катедра), e-mail адрес, с шрифт 11, дясно подравнено.

Образец за форматиране (компресиран Word файл) можете да изтеглите от адрес http://kst-forum.netfirms.com/Spisanie_Obrazec.zip

**Този брой от списанието се отпечатва
със средства от научно-изследователски
проект, финансиран целево от
държавния бюджет.**