MODELING OF THE KNOWLEDGE DYNAMICS OF STUDENTS OR EMPLOYEES

A P SVIRIDOV

Moscow State Social University
107150 Moscow, Losinoostrowskaya str., 24
E- mail: sviridovap@smtp.ru

Keywords: Statistical dynamics of knowledge, intensity of forgetting, knowledge flow, learning flow, open and closed model of knowledge dynamics, computer testing of knowledge.

Abstract. Statistical Theory of Teaching and Learning (STTL) [1-9]. includes statistical dynamics of knowledge [2,4,6-9], the methods of computer testing and diagnosis of knowledge (CTK and CDK) [1,4,5], the algorithms for supervised and unsupervised Learning of Teaching, Learning and Raiting Systems, metrological support of tutorial process. The principles of theory of statistical dynamics of knowledge and tutorial process control have been developed. They include the main ideas and characteristics of dynamics of knowledge in different subtopics, topics and educational disciplines, the intensity of learning and loss of knowledge, intensity of presentation of the educational material and knowledge restoration, the load on the person being trained and other related factors. The methods of the macro- and micro-models of learning and forgeting identification have been considered. On their basis the number of optimum strategies of the training quality's (anti- abnormal training) control, including the error filing when CTK are suggested.

The principles of CTK standardisation theory have been developed. They include the principal ideas, features and methods of analysis and synthesis of CTK plans, with use of quantitative, qualitative, alternative and linguistic sign with errors in identification of correct or wrong answers. The system of plans has been suggested. The system includes: 1) normal, intensifed and short plans; 2) the rules of passage from one plan to another.

1. KNOWLEDGE DYNAMIC PARAMETERS OF THE DIALOGING SYSTEM

The numerical dynamic characteristics of knowledge are determined on the base of training diagrams and retentional curves, which is the temporal dependence of complex index of training quality: 1) the reaction time, 2) the probability how is the work done right or wrong by trainer/operator….

Training diagrams and retention curves can de analyzed on macro- and micro-level [2,4,5,9]. In first case we have got usual exponential macromodels, the dialogue system is made for it's identification.

The definition example of intensity of forgettable λ exponential models decision of the task on the program language: Turbo-PASCAL – from 0,3 to 1,535; C - from 0,8 to 2,07 (per year), the mean time of knowledge keeping by Turbo-PASCAL is more than by C.

Task to solve: 1) Give an example of definition program of minimum element one-dimension massive; 2) Make the procedure of parallelepiped graphic representation.

Transformation micro-models of one structure, as a rule are nonlinear, nonmonotonous, as a rule includes local maximums and minimums and plateau, reflecting evolutional places of perfection some strategies and transition from them to more perfect one. For it's identification it is necessary to use more powerful methods of quantum physics than the theory of probability. There is a new conception among them - complex mark wave function and amplitude of probability. The condition of knowledge "a" and transition to condition "b" we can compare amplitudes of probability A(a) and A(a,b) – complex numbers, module square is probability of condition "a" and transition from this to b: $P(a)=|A(a)|^2$, $P(a,b)=|A(a,b)|^2$. The use of quantum-physic approach will be shown on the process of forgetting. Compare to the process of forgetting with exponential macro-model $Q(t)=Q_0\exp(-\lambda t)$,

where

Q_0- probability to solve right the task in the moment of time t=0,

λ - intensity of forgetting,

$A_1(t)$ and $A_2(t)$ - two amplitudes of probability in aspect:

$A_1(t) =(0,5\ Q_0\exp(-\lambda t))^{1/2}(\sin \omega_1 t + j\sin \omega_2 t)$

Proceedings 17th European Simulation Multiconference
(c) SCS Europe BVBA, 2003 ISBN 3-936150-25-7

and

$A_2(t) = (0,5\ Q_0 \exp(-\lambda t))^{1/2}(\cos \omega_1 t + j\cos \omega_2 t)$

The resulting amplitude of probability A(t) in the moment of time t is defined by sum

$A(t) = A_1(t) + A_2(t)$ or

$A(t) = (0,5\ Q_0\exp(-\lambda t))^{1/2}(\sin \omega_1 t + \cos \omega_1 t + j(\sin \omega_2 t + \cos \omega_2 t))$.

So the probability of solving right of task in the moment of time t is: $Q_1(t) = |A(t)|^2 = 0,5\ Q_0\exp(-\lambda t)(2 + \sin 2\omega_1 t + \sin 2\omega_2 t)$.

Example:

For macro-model of the forgetting process by informative-measuring technique $Q(t)=0,745\exp(-0,018t)$ (t per month) [2,4,9] we have got a micro-model $Q_1(t)=0,373\exp(-0,018t)(2+ \sin 220t+ \sin 499,1t)$.

With meaning ω_1 and ω_2 in accordance with the frequency

$f_l=0,67*10^{-5}$ Hz, $f_2=1,53*10^{-5}$ Hz.

Such a way of micro-models building can be used in training modeling, developing, dynamics of difficult systems, scientific technical and social progress, but wavelike processes and characteristic for them too. That is why we use wider interpretation of "learning".

2. MODELING AND PROGNOSTICATION

For the description knowledge dynamic models the main ideas are used: studying material flow (knowledge flow) L and flow of comprehensibility/rehabilitation of knowledge (learning flow) A. They are determined in the way of putting conforming flows L_i and A_i ($i \in (1,..., N)$) for the separate tasks with their life circles (LC) S_i:

$L=L_1 + L_2 +...+ L_N$, $A=A_1 + A_2 +...+ A_N$ (knowledge and learning flows).

Life cycle (LC) S_i represents a time interval, during which a situation i (or standard activity units SAU)) assimilation is necessary. Subjects flow L_{1i} is formed by time moments t_0^{1i}, t_1^{1i}, t_2^{1i},..., where $t_{k+1}^{1i} > t_k^{1i}$ ($k \geq 0$). A time moment t_0^{1i} corresponds to the moment of the first presentation of the subject in a current situation (SAU) according to the program, and the following time moment t_1^{1i}, t_2^{1i},..., correspond to the forgetting moments.

Knowledge assimilation (recovery) flow A_{2i}, attached to situation i (SAU) is formed by time moments t_0^{2i}, t_1^{2i}, t_2^{2i},..., where $t_{k+1}^{2i} > t_k^{2i}$ ($k \geq 0$). Time moment t_0^{2i} corresponds to assimilation moment of situation i (SAU) after its first presentation, and the following time moments of knowledge recovery after their forgetting.

The analogical way is for gave and done tasks in the other cases of human activity [2,4,9]. All these permits to use dynamic knowledge models in modeling their human activities.

The knowledge dynamic is determined:

1. By the flow of studying material (knowledge flow) X;
2. By the flow of comprehensibility/treatment of knowledge (learning flow) Y;
3. By the number of N task/function in the discipline, which must be done by the student/operator;
4. By the way of learning and treatment of knowledge r;
5. By the number of s channels of knowledge learning and treatment (for example: visual and acoustic channels);
6. The capacity buffer memory m.

The open knowledge models of dynamic are described by 6 symbols X/Y/N/r/s/m (wider symbolism of Kendall). Closed models must be in brackets. There can be M,G,GI,D,E_k insert X and Y. So, Y=M means that time of learning is free and has exponential assessment.

Let's see net dynamic models of knowledge <M/M/N/2/1/(N-1)> and <G/M/N/2/1/(N-1). They are identical of closed one channel stage learning/treatment with N source, which conforms to N task of the specialties (learning disciplines)

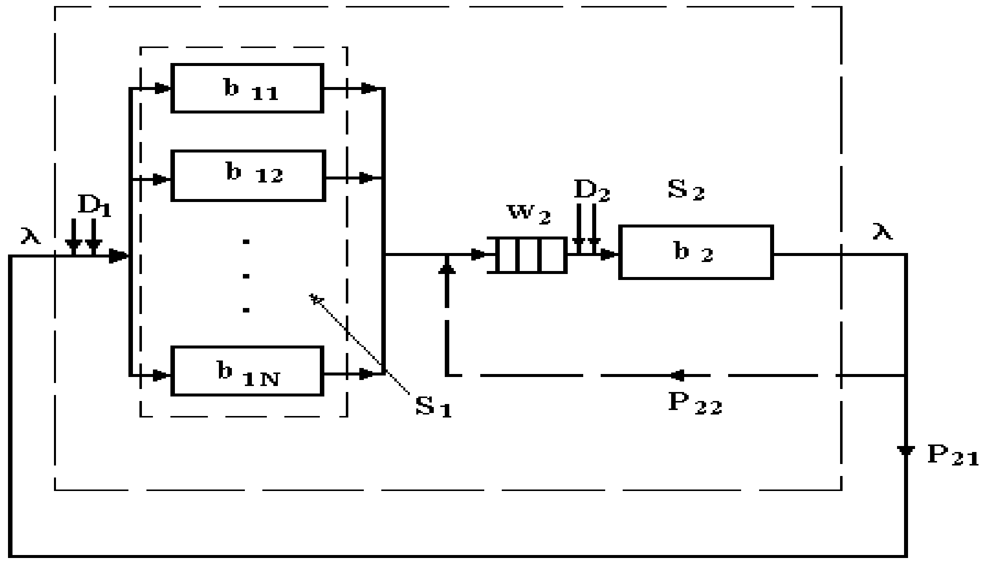

Fig.1. Closed model of Knowledge dynamic

for the trainer, (N-1) place of waiting and to the stage of learning without loosing. The discipline of learning - "First Come First Served" (FCFS). The dynamic model of knowledge (Fig.1) consists of stages of forgetting S_1 and learning S_2, and learning could be done in the time-sharing regime. In this case for learning or reanimation of knowledge in some field of task distinguished time quantum. If this time quantum is enough for learning, than the task with probability p_{21} comes to the stage of forgetting S_1. On the contrary task with the probability p_{22} comes to the stage of learning S_2. Let as mark:

λ – the mean intensity of task forgetting,

μ - the mean intensity of task learning,

b_{1i} (i=1,2,...,N) – the mean time of forgetting time waiting of i position/task/function, $b_2=1/\mu$ - the mean time of learning/reanimation of knowledge by condition/task/function,

D_1 and D_2 – disciplines of forgetting and learning,

w_2 – the buffer capacity for forgetting condition/task/function before the learning stage S_2.

There are more learning channels (visual…) used in the learning process by studying (multimedia) courses, the intensity of learning is higher and intensity of forgetting is lower.

The main characteristics of this model:

The probability that the student is able to

$$p_0 = [\sum_{n=0}^{N} \tfrac{N!}{(N-n)!}(\lambda/\mu)^n]^{-1}$$

solve all N tasks

The average number of tasks waiting time in the system (knowledge reanimation)

$$L_q = N - \tfrac{\lambda}{\mu}(1-p_0)$$

The average waiting time in the waiting line

$$w_q = \frac{L_q}{\lambda\,(N-L)}$$

The average number of forgetting tasks

$L=L_q+(1-p_0)$

The average waiting time in the system (reanimation)

$w=w_q+1/\mu$

The probability of forgetting n tasks

$$p_n = \frac{N!}{(N-n)!}(\frac{\lambda}{\mu})^n p_0,\ n = 0,1,...,N$$

For the programming language C we have got: N=51, λ=1,05 [1/per year]=1,22 10^{-4} [1/per hour], μ=1, $p_0 \cong 0,994$, $p_1 \cong 0,006$, $L_q \cong 1,56$, w_q=260 hour, $L \cong 1,566$, $w \cong 261$ hour.

Mounted: optimum sequence of setting relative priorities in reanimation of forgetting tasks/functions are determined by relation c_i/b_{2i}, where c_i – is the fine/waste for the lack of knowledge i position/tasks/functions (for disability of making i tasks or functions). Close net dynamic models of knowledge might be used as mathematical model in case of task solving ensuring **guaranteed quality** of professional training for N tasks solved or **guaranteed quality** of student's training for the final number of training disciplines.

3. OPTIMIZATION OF QUALITY MANAGEMENT

The forgetting intensity might be use for optimal planing of lerninig material planing (LMP) [2,4,9]. A number of optimal strategies of preparation quality control are developed using:

- intensity of assimilation and forgetting;
- economic factors (education and control costs, losses due to the error actions, possible benefits, etc.);
- under-estimate or over-estimate risks on the CTK.

Let us analyse two strategies:

Strictly periodical organization strategy of LMP.

Taking into account knowledge under-estimate (α) and over-estimate (β) risks and correctness of exponential time distribution, a learner (operator) professional availability ratio is determined as follows [4,9]:

$$K(a)=\frac{(1-\beta)\,(1-\exp(-\lambda a))}{\lambda(a+t_k)(1-\beta\exp(-\lambda a))+t_a(1-\beta)[1-(1-\alpha)\exp(-\lambda a)]},$$

where λ - intensity of forgetting, t_a - average time of knowledge recovery, t_k - average time of knowledge control, and a - interval between knowledge tests and repetitions (recovery). For an operator who has λ=0,1 month^{-1}, t_k =0,5 h, t_a = 3 h, α=0,35 and β=0,4, an optimal interval between tests and anti-failure trains will be: a_0 = 110 h. Note that maximum value of $K(a_0)$ is 0,97.

For trainers we have got: the intensity of forgetting λ_1=0.05 и λ_2=0.40 1/month, average time of knowledge reanimation t_e=3 hour, average time of knowledge control t_k=0.5 hour, the underrate or overrate risk on the computer testing of knowledge (CTK, computer-based test) α=0,35 and β =0,40. Optimal time intervals between repeating (anti damage training), guaranteeing the maximum coefficient of professional training, in case of periodical management strategy of training for 1 and 2 trainers/operators is T_1=160 hour, T_2=50 hour. In case of LMP (anti damage training) using through the intervals, the coefficient of workers professional training is lower.

Strictly periodical strategy of LMP with two checking types.

N=(1,2,...,n) – multitude position of tasks, functions which must be learn by students/users. They should pass tests for professional training guarantee complete and part knowledge checking are used with further knowledge maintaining on forgetting ideas and tasks. If the test is global - whole knowledge is controlled and if the test is partial just part of knowledge is controlled**A** (**A**$\subset$**N**).

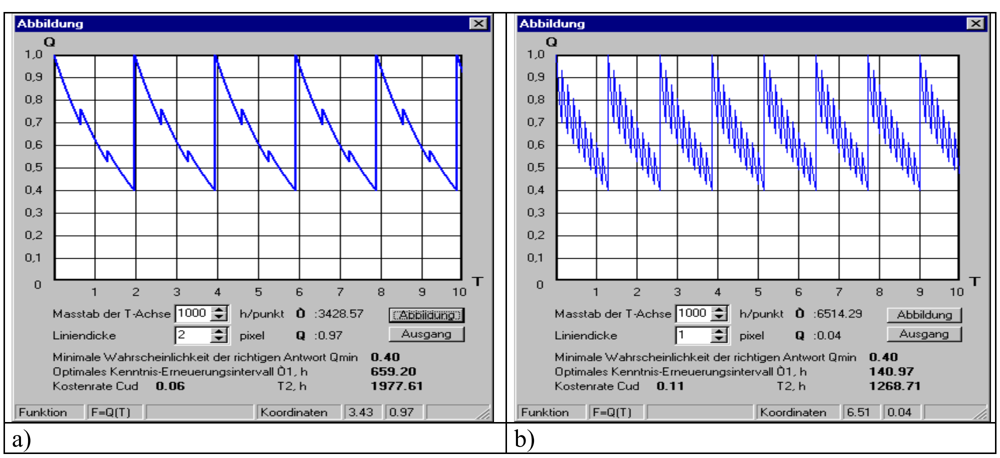

Fig 2. The probability dependence of correct tasks by students/users.

λ_1 and λ_2 – are the forgetting intensity of tasks **A** and **N\A,** c_1 and c_2 (c_2>c_1) – the expeditures for test organization and reanimation of forgetting knowledge **A** and **N**, a_1 and a_2 – intervals betveen global and partial tests, in this case a_2=ra_1 (r>1). The optimization problem is definition of meaning

a_{10} and a_{20}, the average expenditures are minimum, and probability of correct tasks $Q(a_2\text{-}0)$ is not below of some Q_0. On the Fig.2 are dependences Q(t) for Q_0=0,4, c_1=5, c_2=100: a) λ_1=0,000139 [1/hour], λ_2=0,000417 [1/hour]; b) λ_1=0,002 [1/hour], λ_2=0,0005 [1/hour]. The answer is: a) a_{10}=659,20 hour, a_{20}=1977,61 hour; b) a_{10}=140,97 hour, a_{20}=1268,71 hour.

The strategies have realization in dialogue system. These models might be used in integration of business and knowledge management.

4. STANDARDIZED METHODS OF CTK

Statistic tool CTK is a plan determined by a number of tasks/questions that taken from (potential) general combination for the learner, by volumes of the tasks and by conditions of giving different grades. Such plans are widely used for control and preparation quality certification by Universities and companies.

Main characteristics of CTK plans

1. Operating characteristics OC,
2. Risks of underestimation and overestimation of knowledge α and β (error of the first kind, α-error,
   error of the second kind, β-error),
3. Realization of psychometric function of teacher,
4. ASN=average sample number,...

Synthesis of equivalent and ε-equivalent single sampling plan inspection of computer-based knowledge test while understanding the reality of answers with and without errors (for example in choosing way of entering the answers=Multiple-Choice-Input) can be done: 1) by two dots OC (P_1, 1-α) and (P_2, β), where P_1 - max. part of misunderstood questions when "credit" is given, P_2 - min. part of misunderstood questions when "no credit" is given, 2) by dot of indifferent OC (IQL=indifferent quality level) and curve in it (plans of ($P_{0,5}$,h) kind); 3) by meaning P_1 and α or P_2 and β; 4) by giving psychometric function of teacher/expert; 5) by economic data. It is better to use Larson-nomogram for graphic synthesis.

Example.
Given: P_1=0,2, α=0,3, P_2=0,3, β=0,35. On the basis of Larson-nomogram of the cumulative binomial distribution we have: n=12, c=3, i.e. learner is offered 12 questions (tasks) "credit" is given if max. number of errors is not more than 3 (single sampling plan (12,3)).
Let it use choosing way of entering with average number of answers for choosing S=4, than on the basis of Larson nomogram we have: n=18, c=3 (plan (18,3)). Plans (12,3) and (18,3) are equivalent from the point of view of risks of underestimation and overestimation of knowledge.

Plans of knowledge control composition

Let's enter the following levels of control: 1) task, 2) section (for example lab work), 3) theme (part of activity), 4) learning discipline (activity), 5) preparation quality on specialty in University, 6) organization/institute, 7) (higher) education. Correlation between elements of different levels of correction -"**consists of**". Number of level of correction - **difficulty degree of control system, diagnostics and certification.** Let us see the way of determination of OC composition of control plans of 2 neighbor levels.
Let the correction on one level be done according to the plan with OC L_1*(Q). Such level can be: a) knowledge control of separate student on different sectors of learning discipline, b) knowledge control of students for the whole learning discipline or specialty. On the next level the decision is being made according to the second plan with OC L_2*(Q), i.e.:
a) about understanding of the whole learning discipline by the student,
b) about the preparation quality on specialty or learning discipline in organization/university.
The result OC of control plan on the second level looks like:

$$L_{res}^*(Q)=L_2^*[L_1^*(Q)].$$

Example.
For two-mark CCK of students of some specialty is used one step plan (10,2). The conditions of preparation quality certification on specialty on the results of correcting separate students according to the plan (10,2): P_1=0,1, underestimation risk α=0,2; P_2=0,3, overestimation risk β=0,35. It is necessary to synthesize one-step plan of preparation quality certification on specialty in the university.

Solution.
We have:
P_1^*=0,1*0,65+0,9*0,2=0,245,
P_2^*=0,3*0,65+0,7*0,2=0,335.

On the basis of Larson-nomogram we get: n=45, c=11. This means: one should choose 45 students from general combination and correct them according to the plan (10,2). If the number of unprepared students/operators is not more than 11, than the preparation quality on the specialty in organization/university is positive.

Ways of analysis and synthesis of computer-based knowledge-tests for various tests and examination layers (especially for individual students, employees and courses of studies) are realized in dialog system.

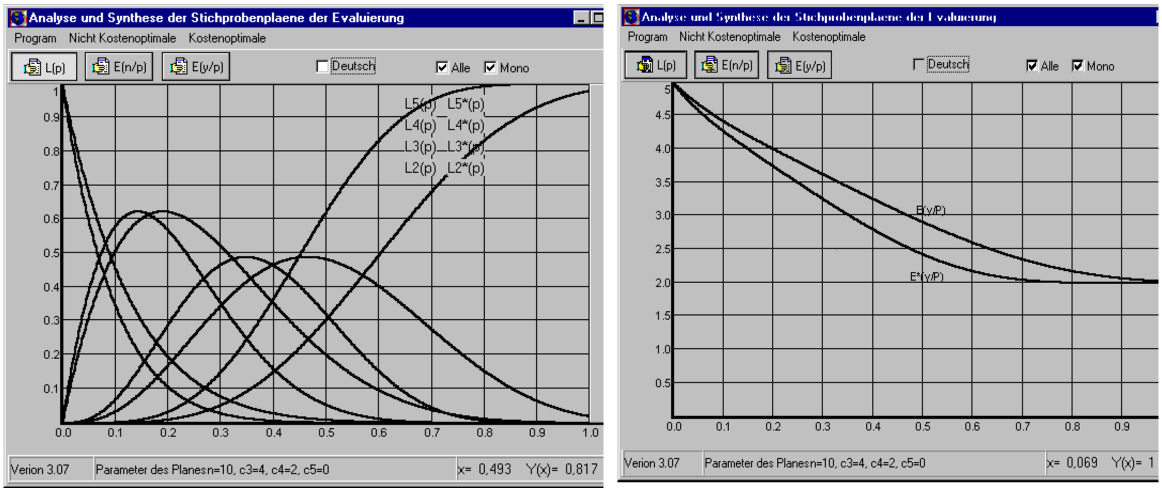

Fig. 3: a - OC $L_i^*(P)$ and $L_i(P)$, i=2,3,4,5; b – realization $E^*(y/P)$ and $E(y/P)$ psychometric function of teacher/expert

Let us see the examples of its using.

On Fig. 3a and 3b are shown:

- $L_i^*(P)$, i=2,3,4,5 - OC of one step plan $(10,0,2,4)^a$ of four mark CCK without mistakes in determining the reality of the answer (writing or oral answers),
- $L_i(P)$, i=2,3,4,5 - OC of the same plan $(10,0,2,4)^a$ in choosing way of entering the answers, when approximately 4 possible answers are given for each question; these OC are moved right to OC $L_i^*(P)$;
- $E^*(y/P)$ - realization of psychometric function of teacher/expert (dependence of average mark y on possibility of wrong answer P) with the help of intellectual learning or supportive system that realizes plan $(10,0,2,4)^a$ according to writing or oral answers (low curve),
- $E(y/P)$ - realization of psychometric function of teacher/expert with the help of intellectual learning or supportive system of the same plan $(10,0,2,4)^a$ with choosing the way of entering the answers when approximately 4 possible answers are given for each question (high curve).

REFERENCES

[1] Sviridov A.P. - „Introduction to the Statistical Theory on Teaching and Learning. Part I. Sampling Plans for Computer-Based Knowledge Tests“. (Wwedenie w statisticeskuju teoriju obucenija. Cast 1. Standartisirowannye metody kontrolja snanij). - Moscow: Published by "Moscow Power Energetics Institute", 1974.- 134 P.

[2] Sviridov A.P. – „Introduction to the Statistical Theory on Teaching and Learning. Part II. Statistical Knowledge Dynamics“. (Wwedenie w statisticeskuju teoriju obucenija. Cast 2. Elementy statisticeskoj dinamiki snanij). - Moscow: Published by "Moscow Power Energetics Institute", 1974.- 152 P.

[3] Sviridov A.P. – „Supervised and Unsupervised Learning of Teaching, Learning and Raiting Systems“. (Obucenie i samoobucenie obucajuschich i kontrolirujuschich maschin). - Moscow: Published by "Moscow Power Energetics Institute", 1976.- 182 P.

[4] Sviridov A.P. – „Fundamentals of the Statistical Theory on Teaching and Learning“. (Wwedenie w statisticeskuju teoriju obucenija i kontrolja snanij). - Moscow: Published by "Wysschaja schkola" (Higher school), 1981.- 262 P.

[5] Swiridow A.P. – “Statistische Lehrtheorie und Humanisierung von computerunterstuetzten Lehrsystemen”. -IBM-Hochschul-Kongress. Dresden 30.9-2.10.92. Proceedings.

[6] A. Swiridow, I. Schalobina. – “Kenntnis-Dynamik ueber ein Lehrfach”. -Ilmenau: 42. Internationales Wiss. Kolloquium an der TU Ilmenau, 1997, Bd.1, pp. 401-406

[7] Swiridow, D. Slesarew, I. Schalobina. “Simulation von Kommunikationssystemen und –netzen als sozio-technische Systeme”. In: 43. Intern. Wiss. Koll. An der TU Ilmenau.-Ilmenau: TU Ilmenau, Bd.1, 1998.- pp. 174-180

[8] A. Swiridow, D. Reschke, N. Slesarewa. - “Modellierung der Kenntnis-Dynamik”. In: 43. Intern. Wiss. Koll. an der TU Ilmenau.-Ilmenau: TU Ilmenau, Bd.1, 1998.- pp. 282-189

[9] R. Suesse, A. Swiridow. – “Statistical Knowledge Dynamics”. (Statistische Kenntnis-Dynamik).-Ilmenau: Wissenschaftsverlag, 1998.- 256 P.