acmbul 93

COMPUTER APPLICATIONS

Eastern European Regional Programming Contest

acmbul 93

PROCEEDINGS

INTERNATIONAL CONFERENCE

16-20 October, International House of Scientists
St.Konstantine Resort, Varna, Bulgaria

COMPUTER APPLICATIONS'94
16-20 October 1993, Varna, Bulgaria
Organized by ACMBUL

10

Simulating the Software Market

Avram Eskenazi, Rossen Rashev
Institute of Mathematics, Sofia
Acad.Bonchev str.bl.8, 1113 Sofia
phone ++359 2 7132873 E-mail: BANMAT@BGEARN.BITNET

Abstract

A simulation model of the software market proposed by Swann is considered. Refinements and enhancements of this model are suggested. A program implementing the enhanced model has been realized and experiments have been performed with. Possible applications are discussed.

1. Introduction.

Modelling the software market might have a positive effect on software companies' policy. It might also influence the selection of new products development projects - as is well known an important factor in this respect is the relationship between market structure and innovation [1]. Simulation technique is one of the promising ways to build such models. We have based our study on Swann results [2], [3], [4]. That explains the structure of this paper - after a brief description and analysis of Swann's model we propose several refinements or enhancements aiming at a more adequate representation of the real market. At the end we briefly discuss possible applications of the model.

2. Swann's software market model.

According to [5], [6], [8], [9] users' choice is influenced by both the intrinsic quality of products and the network externalities these products offer. Though difficult, the problem of quality is relatively clear and might be solved in one or other way. The network around the product is related to the benefits the user might acquire - number of (other) users, "add-on" products, courses, factors reducing the risk of buying an obsolete product. It is assumed users have the following "indifference" function:

$$(1)\quad U(i) = a(i) + b(i)*m(j)*N(j) + c(i)*Q(j)$$

where $Q(j)$ represents the intrinsic quality of product j, $N(j)$ is the installed base (number of installed copies) of product j. $m(j)$ is an additional differentiating factor used to simulate the "intensity" of network externalities per unit installed base. $m(j)*N(j)$ represents the network. $a(i)$, $b(i)$ and $c(i)$ are parameters of user's i indifference function.

Normalizing (1) we obtain:

$$(2) \qquad U(i) = N(j) + c(i)*m(j)*Q(j)$$

Note, that the only parameter depending on i is $c(i)$. That means that users are supposed to have the same assessment of $N(j)$ and $Q(j)$ for each product j but differ in their relative valuation of intrinsic quality to network externalities. Thus each user i is assigned a particular $c(i)$. The set of all $c(i)$ defines the types of users on the market. We presume that all $c(i)$ belong to the interval $[0, Cmax)$ and that their distribution is uniform.

Another factor represented by the model is the pre- announcement. As is well known this marketing technique is used to delay users' decisions. A user may buy an available product, but he also may delay his purchase waiting for one of the products pre-announced. In the latter case (2) is modified to:

$$(3) \qquad U(i) = (m(j)*N(j) + c(i)*V(j))/(1+r)^{t},$$

where t is the number of waiting periods until the product j appears on the market and r is the so-called 'discount-rate' (DR) which decreases $U(i)$. According to Swann DR is constant for all users. Users are supposed to select between buying immediately and waiting for pre-announced products after comparing the results of (2) and (3). Obviously, this comparison is typically intuitive or at least - more intuitive than analytic.

There is one basic assumption in the model - that users of a given type program products appear in a steady and increased stream, defined by an S-shaped logistic growth curve:

$$(4) \qquad n(t) = nmax * \frac{exp(f+gt)}{1 + exp(f+gt)},$$

where $n(t)$ is the number of new users arriving at time t, g is a growth rate parameter, $nmax$ is the saturation level of new consumers, and f is an intercept parameter.

The model simulates the market by determining for each period t the number of arriving users. These users are distributed between availab[illegible] and pre-announced products, by using (2) and (3). This distribution reflects on $N(j)$ of each product j and influences the users' choice at [illegible] period. This dynamics gives the model its rather [illegible] properties:small changes in the distribution of tastes, in the [illegible] of demand, and in the introduction dates can give very different re[illegible]

3. Further development of the model.

We developed a program implementing the Swann model. Observing the results obtained, as well as those given in [2], we found two essential deviations from the real processes.

a) The simulations do not show all products purchased in a given period, but only two or three among which all purchases are distributed.

b) Some products which are present on the market do not appear at all for the whole simulation period.

In order to overcome these drawbacks and in an attempt to expand the scope of the model we propose the following changes:

3.1. Representation of users' tastes distribution.

The Swann model assumes that users' tastes are represented by the $c(i)$ parameter (2), which may vary in the interval $[0, Cmax)$. Users are uniformly distributed into 100 types. Each type i is assigned $c(i) = i.Cmax/100$. (By the way, Swann is establishing the value of $Cmax$ after several simulation tries). We think the hypothesis of the uniform distribution is not enough well-grounded and we suggest two ways to find out a more adequate distribution.

Let's denote by $W(i)$ the number of users of type (group) i divided by the total number of users. Wi calculate $W(i)$ by means of following procedure:

- we select a period from the time interval we are modelling, such that there are no pre-announced product;

- for each product j we collect data about:

  = the number of copies $S(j)$ sold up to the beginning of the period in consideration;

  = the number of copies $dS(j)$ sold during this period;

- we sort the products on $S(j)$ and $Q(j)$ in descending order (remember that $Q(j)$ is the intrinsic quality estimation value);

- taking the indifference function we work out the equation

$$c*Q(1) + m(j)*N(1) = c*Q(2) + m(j)* N(2),$$

where the indexes 1 and 2 respectively refer to products 1 and 2 ($j=1,2$) assuming that $N(j) = S(j)$;

- by solving this equation we find c ($c = C(1)$)

All users from the user groups i corresponding to the value $c(i)$ in the interval $[0, C(1))$, have chosen product $j=1$, because it possesses the highest value of quality and network externalities. For each user group i we define $W(i) = dS(1)/(nu(1)*\Sigma dS(j))$. $nu(1)$ is the number of user groups in the interval $[0, C(1))$ ($nu(1) = C(1)*(Cmax/100)$).

- similarly we calculate $C(2)$, $C(3)$, ... while it is possible.

To each user group i, corresponding to a value $C(i)$ from the interval $[C(k), C(k+1))$, we assign $W(i)=dS(k)/(nu(k+1)*\Sigma dS(j))$. $nu(k)$ is the number of user groups in the interval $[C(k), C(k+1))$ $(nu(k+1)=(C(k+1)-C(k))*(Cmax/100))$.

We assign $W(i)=dS(l)/(nu(l)*\Sigma dS(j))$ to the user groups in the interval $[C(l), Cmax)$, l being the last product number. Here $nu(l)$ is the number of user groups in the interval $[C(l), Cmax)$ $(nu(l)=(Cmax-C(l))*(Cmax/100))$.

Please note that unsolvable siltations may occur when the indifference function used doesn't correspond to reality. They may be treated as a signal for inadequate data, incorrect indifference function form or too small value of $Cmax$.

The second method consists in selecting the most appropriate distribution function after several simulations with different distributions.

3.2. Modelling the quality. The model is very sensitive concerning the products' quality. Small changes in the users' quality estimation immediately change the end simulation results. But there are at least two problems - quality estimation is a very heavy procedure and it is essential to represent the users' view on particular product quality. An additional requirement should be the knowledge of the quality rank of each product, as well as its precise distribution.

To make the problem more difficult (but the model more adequate) we have to ask two more questions:

- do all users have the same quality estimation for a given product?
- does this estimation remain constant over the time?

The answer to both questions is "no". The vertical and horizontal differentiation is likely to solve the first problem to some extent. The second 'no' is caused at least by the change of prices (as is well known the majority of quality models include the price as a fundamental quality factor). That is why we modify Swann's assumption for $Q(j)$ being constant for each product j. If an investigation among users were possible, then it would have shown a differentiation of quality estimation, and more precisely - a different estimation of the importance of various quality factors (characteristics). Consequently, $n(t)$ for each period of time should be broken down into several groups depending on the users' views on quality (and, certainly, not on the quality estimation of a particular product).

3.3. Differentiation of the inherited proportion. As already mentioned a new version or an upgrade of a product takes over a given percentage of the source version's network. The analysis has shown that this percentage may vary form one product to another. That is why we assume that different percentages might be given taking into account:

- how far has been kept the user interface;
- to what extent is it possible to use with the new version the data accumulated while using the previous one;
- will a new training be necessary for the new version and how long;
- will these additional efforts be commensurable with the benefits expected from the new version.

There is another marketing technique, which also should be represented by using this percentage. This is the so called "competitive upgrade" - users of any product are offered the chance to purchase a new version of a given product at the cost of the upgrade only (regardless of the fact that they are not licensed users of the previous version). One possibility to represent this phenomenon is to add to the new (diverting) product's network a percentage of the network of each of the products left by the users.

3.4. Differentiation of the pre-announcement effect (DR). The representation of the pre-announcement effect makes the model more adequate. This effect depends on the quality of the pre-announced product, on the network around it and on the time to its appearance on the market. All of them are directly represented in (3). Other factors are represented in the function through the parameter r. The latter depends on:

a) The advertisement realized by the manufacturer (the pre-announcement itself is a kind of advertisement). The stronger the advertisement the more users will get acquainted with the product, and hence - more users would be prepared to wait for the appearance of the product on the market. Therefore a lower r should correspond to a stronger advertisement campaign (remember that r is in the denominator).

b) Another factor influencing r is the image of the company making the pre-announcement. The better the image the lower the value of r.

c) Another factor determining r is the ability of the company to keep its promise. If the date pre-announced is not observed it is possible that part of the potential customers would be lost. Therefore, an unobserved term increases the value of r.

It is possible to analytically represent r as a function of these three factors.

3.5. A more precise representation of the dynamics of the market. It might be realized through an interaction between the end user and the program implementing the model.

The model described in 2. does not take into account a few changes that might happen on the market:

- change of the price of one or more products considered;

- change of a pre-announced date for a product to appear;

- change of the value of r due to modifications in the advertisement campaign of a manufacturer;

- change of some $m(j)$ due to a change of the relation (number of courses and add-ons against total number of installations of a given product).

The program we have developed allows the user to describe the "situation" at the beginning of each time period. More precisely, parameter values which may be entered are: quality $Q(j)$, introduction date, take over (inheritance) percentage, network factor $m(j)$, and pre-announcement factor r.

4. Applications.

The model has a rather broad area of application. It might help a company when taking decision concerning the development or marketing policy of the company.

4.1. Software product life cycle.

Without going into details we can say that at each phase (except the design and programming) the simulation model can contribute to the establishing of a better development policy.

4.2. The model can contribute to a better marketing policy. As stated by Philip Kotler [10] a fundamental tool in this respect is the marketing information system. One of the four components of such a system is the subsystem for marketing analysis and the system in consideration might fulfill this function.

When developing the marketing mix a decision has to be taken concerning each component. That implies a good and reliable understanding of the way the consumer (user) will react to each action. An effective and cheap way to replace experiments in real conditions is to make use of the simulation model.

Conclusion.

We expect that the refinements and expansions we proposed in this work make the Swann's model more adequate and more usable. The program implementing it, is in use and shows promising results. A lack of sufficient real and reliable data from the Bulgarian software markets poses some problems which we expect to overcome in the near future.

This work has been partly supported by the Ministry of Education and Science under contract I24.

References

1. Ali,A., M.U.Kalwani, D.Kovenock. Selecting product Development Projects, Management Science, vol.39(3), pp.255-276.

2. Swann,P., H. Lamaison, "Vertical Product Differentiation, Network Externalities and Market-Defined Standards: Simulation of the PC Spreadsheet Software Market", Brunel University, CR ICT Discussion Paper, (December 1990)

3. Swann,P., H. Lamaison, "Horizontal Product Differentiation, Network Externalities and Market-Defined Standards: Simulation of the PC Word-processor Software Market", Brunel University, CRICT Discussion Paper, (December 1990)

4. Swann,P., M. Shurmer, "An Analyses the Process Generating De Facto Standards in the PC Spreadsheet Software Market", Unpublished Paper, Center for Business Strategy London Business School, (October 1992)

6. David,D.,A., "Clio and the Economics of QWERTY", American Economic Review, (May 1985)

7. Farrell,J., G. Saloner, "Standardization, Compatibility and Innovation", Rand Journal of Economics, (Spring 1985)

8. Farrell,J., G. Saloner, "Installed Base and Compatibility: Innovation, Product Pre-announcements and Predation", American Economic Review, (December 1986)

9. Katz,M., C. Shapiro, "Network Externalities, Competition and Compatibility", American Economic Review, (June 1985)

10. Kotler,P., 'Marketing Essentials', Prentice-Hall, Inc., Englewood Cliffs, 1984