Modelling The Dynamics, Complexity And Evolution Of GRID Users

R. H. El-Abiad & L. E. Sacks

Department of Electronic and Electrical Engineering
University College London

Abstract: The Grid platform, an ecosystem whereby nonlinear dynamics, complexity and chaotic behavior exhibit themselves in the form of convoluted interactions between the users' species, and the socio-economic environment.
This paper studies this behavior with a simple agent-based model using NetLogo simulation tool and delve into the evolution of heavy users in the Grid community.

1 Introduction

The Grid is the global computing infrastructure that sustains the distributed high-performance computing and data-intensive applications dominating the realm of science nowadays. It allows the discovery, allocation, management and monitoring of resources regardless of their geographical position. Grids integrate networks, computing clusters, data, storage devices, processing power and high-speed memories to provide a virtual global seamless integrated platform for high-performance computing [3]. Nowadays the experiments led by high-energy and particle physicists, astronomy-researchers, combinatorial chemists, molecular biologists, computer scientists and engineers pioneer the present Grid topology with big needs for processing powers, data storage, analysis, modelling, visualization, and simulation tools. Nonetheless the Grid will soon be the infrastructure not only for e-science but also for e-business, e-government, e-society and e-life with the emergence of fields like social-science, anthropology, archaeology, archiving, business, economics, health and environmental sciences.

The future evolution of the Grid is neither linear, nor easily predictable, nor even anticipated. A gamut of factors and conditions interact to enmesh the complexity behind the heterogeneous, autonomous, competing and cooperating user groups cohabiting in the Grid system. In order to capture the complexities, interactions and outcomes in this chaotic environment, an agent-based model is designed and implemented using the NetLogo simulation tool, developed by the Centre for Connected Learning and Computer-Based Modelling in Northwestern University, USA.

Initially the microscopic settings of the model are easily interpreted and defined by a set of comprehensible clear rules. However upon running the model, the outcome and long-term shape of the system change drastically and the macroscopic view of that latter is not easily understood. Such is the exact behavior and definition of a complex system whereby chaos and non-linearity dominate the dynamics of the overall system.

2 The Model

The approach adopted in this model is to emulate the Grid community as an ecosystem where the user groups and software and hardware resources form the different species inhabiting the structure grouped by the common traits they share. The rudimentary rules that dictate the basic behavior of the overall organism are the essential survival mechanisms and strategies that living entities reveal socially. Some of these interactions modelled in our system are competition, cooperation, adaptation and on top of all feeding process on other trophic levels. This basic sequence of activities can be depicted by the simple food chain that links the species of the environment through the energy flow e.g. Users feeding on technology that feed on resources. However the connections between these species are much more convoluted in real world and the genera are more complexly entangled leading to a web of resource flow and species' dependence.

Users → Technology → Resources

The Grid system is drawn as an ecosystem having the previous basic food chain describing its trophic levels. It consists of its users groups, its technology platforms (software), and finally its resources devices (hardware). This food chain hides more complexity and convolution among the species that coexist in several trophic levels and that feed on one another in a web-like rather than sequential chain-like way. These three species represent our agents in the model.

The first agents in the model represent the user groups who coexist in the Grid community nowadays. They are characterized by heterogeneity, based on the difference in the weights of the jobs they run over the Grid; autonomy, depicted in their independent behavior; competition and cooperation, which they exhibit among themselves for survival purposes; adaptation, encompassed by their ability to adapt, and learn from the changes and actions of themselves and their environment [1].

The second ones are the technology platforms who abide to Moore's law that states roughly that the cost of a given amount of computing power halves every 18 months [4] and thus the abundance of platforms and advances in technology doubles every 18 months.

The third and final agents are the resources from computing and processing powers to memory and storage devices aggregated in clusters of high-performance computing machines.

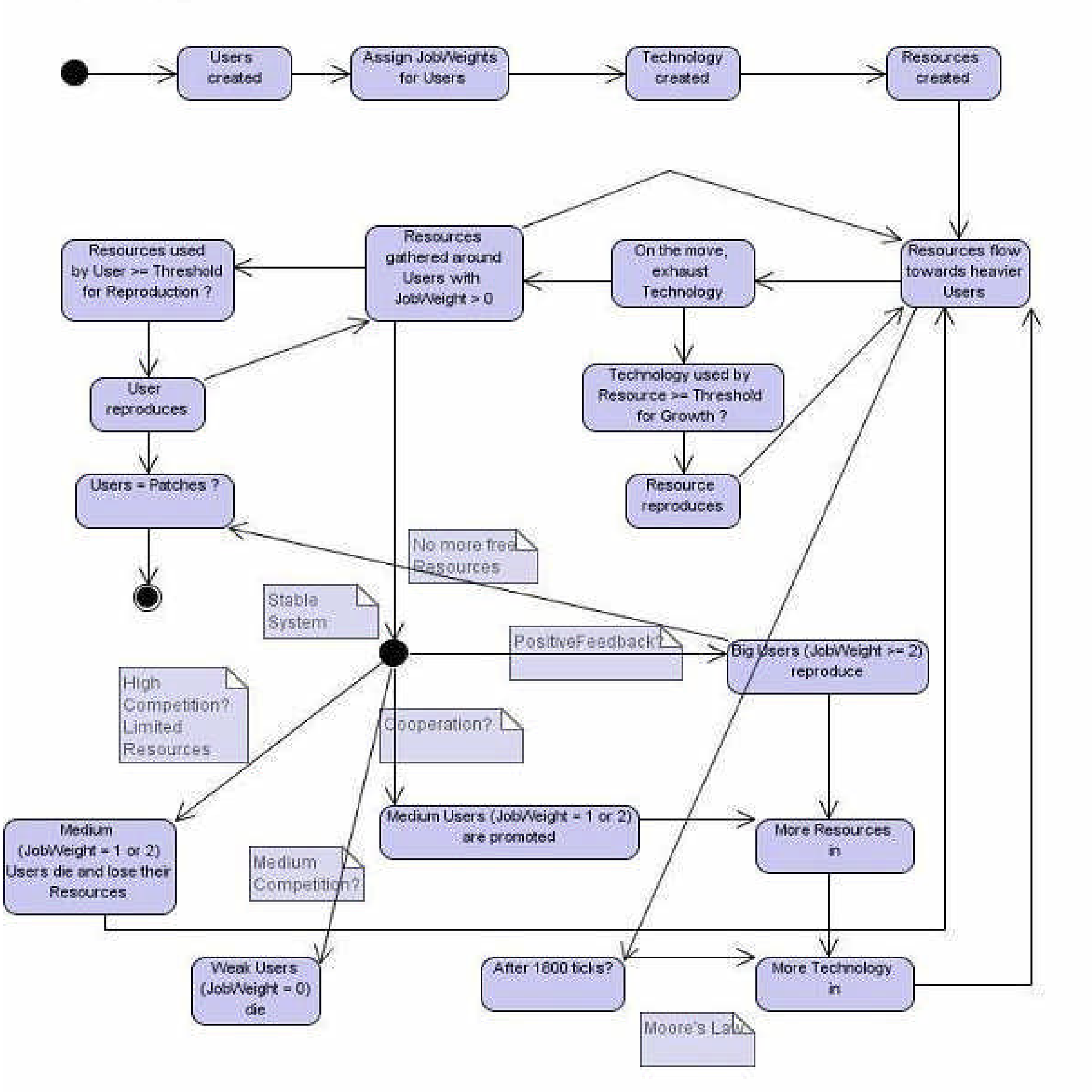


Fig.1 State Flow Diagram of the Model

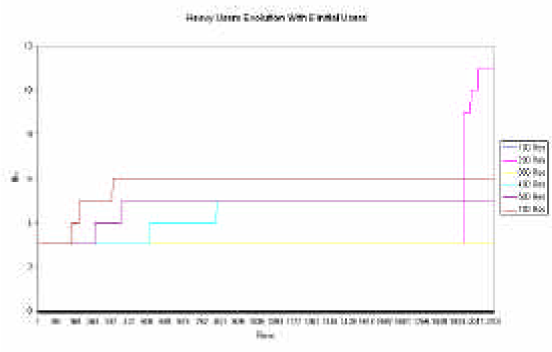
Competition, cooperation and distribution of the genera affect the abundance of each specie. The improvement of technology and its exploitation by the manufacturers and thus the resources produced augment the abundance of these resources in the economy when a certain threshold for growth is met; Big users hatch when a certain threshold for reproduction is crossed, meaning the creation of new niches in the community based on the increasing returns and positive feedback that dominate the economy [2]; Moore's law predicting the ever-increase of technology [4] and Gilder's law specializing in the advances of optical networks [3] especially helpful to Grid computing; plus other socio-economic factors establish the initial conditions of the system, which are caught by the final outcome and behavior of the ecosystem.

The rules by which the model evolve are described by the above state diagram [Fig.1]. It captures the essential conditions and states that the system passes through.

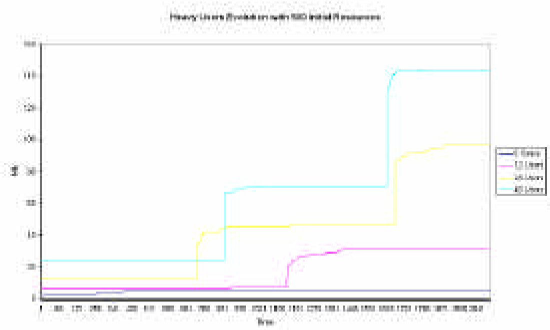
3 Results and Analysis

The aspect in study is the evolution of heavy users in the Grid community. We mean by heavy users the user groups that submit and run hefty jobs on the Grid infrastructure, requiring a lot of computing and processor power or/and big storage devices that could not be easily located locally.

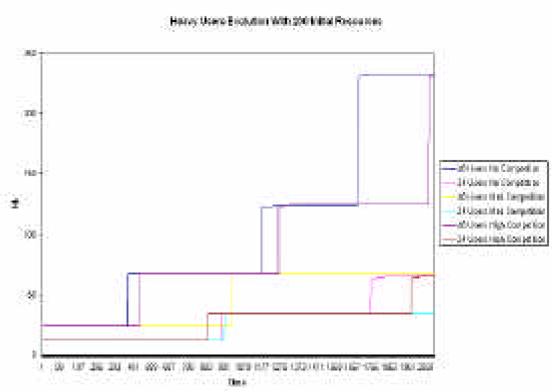
After running the model with different parameter combinations, graphs were generated that illuminated the effects of competition, cooperation, abundance of technology, abundance of initial adopters and abundance of initial resources available in the system on the final growth of heavy users.



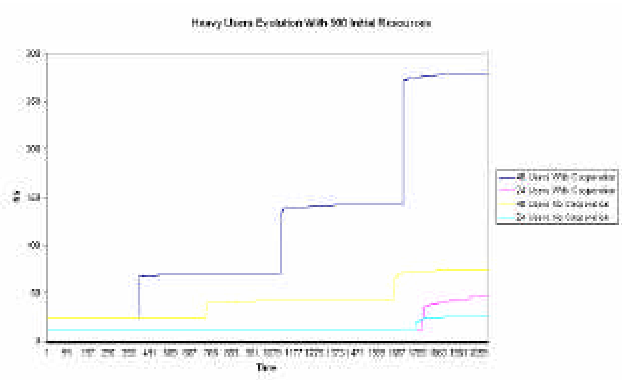
Graph 1. Heavy Users Evolution with 6 Initial Users Varying Initial Resources Abundance



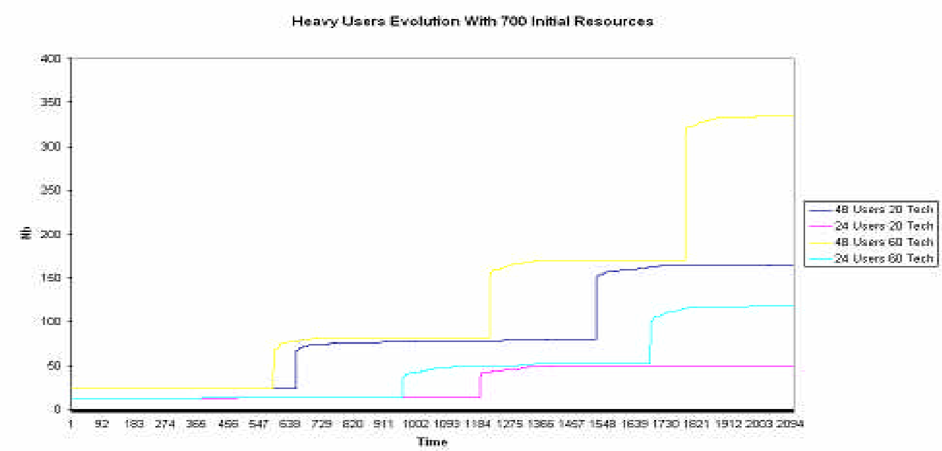
Graph 2. Heavy Users Evolution with 500 Initial Resources Varying Initial Users No



Graph 3. Heavy Users Evolution with 200 Initial Resources with No/Med/High Competition



Graph 4. Heavy Users Evolution with 500 Initial Resources with/without Cooperation

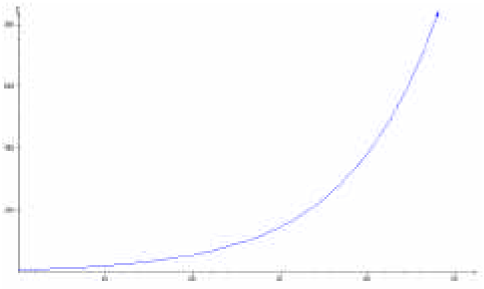


Graph 5. Heavy Users Evolution with 700 Initial Resources varying Technology Abundance

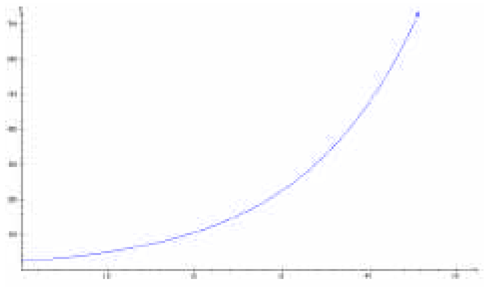
We noticed that the final number of heavy users is not only dependant on the input number of initial adopters or resources available, in fact the biotic settings i.e. the environmental conditions set at the time play a major role in the final outcome of the system [graphs 1-5]. Mathematically speaking the system described is not linear; rather it is nonlinear where a small change in the initial parameter space or any noise appearance in the environment changes drastically the output and behavior of the model.

The combination and analysis of further graphs and data led to the formulation of a mathematical equation that describes this pattern. Actually we measured the final number of heavy users achieved in each experiment, for different values of initial users in the system, and for different amount of resources available in the model, assuming a typical sample of a Grid environment where competition and cooperation between the users are high and technology is available and increased based on Moore's law every 18 time clicks. Positive feedback in the economy [2] is clearly established in our model, especially symbolized by the poll-in of more users, the more initial users we have.

We observed that the growth of heavy users at the end of the simulation run is exponential relative to the initial amount of users present in the system. Speculating that it should be of the form $y = a\,e^{bx}$, in order to find b which describes the rate of this exponential growth, and a the scale factor, we took the natural logarithm of both sides and ended up then with a linear equation of the form $lny = lna + bx$ i.e. $Y = A + bx$. After finding the best linear fit of the values for a log-log scale, we basically found the exponential equation describing the behavior of this system.



Graph 6. Growth with 100 Initial Resources



Graph 7. Growt h with 700 Initial Resources

The equation describing the growth is in case of the 100 Initial Resources: $y = 8.166\ e^{0.0964x}$ and for the 700 Initial Resources, it is: $y = 23.807\ e^{0.075x}$. We continued to generate the equation describing each system behavior with different initial parameters; it always ended up with an exponential form. The increasing returns implemented in the economy of the model pour more resources and users after the system goes into stability, i.e. all available resources are used; and then the process of migration and flow of resources towards heavy users and the augmentation of technology is repeated again, which might yield to the very first initiation of a fractal-like shape of the overall progression curve of the system, as fractals usually require an infinitely many orders of magnitude of power-law scaling in a pure mathematical sense [6].

Also, to involve the abundance of resources in the equation, we then observed the evolution of b (i.e. the slope or rate of growth of the curve) against the abundance of the initial number of resources available and we conclude that wherever the more resources are initiated in the system, the larger is the amount of final heavy users.

4 Conclusions

The study of the Grid as ecosystem and the species interactions among the environment revealed a possible exponential increase in the heavy users in the community. This conclusion is the result of an evolution studied with some predefined initial assumptions and conditions of the system; however the system is sensible to initial dependence [5]. Any change in the parameter space will evolve into a different overall behavior winding up with chaotic complex dynamics of the system [5].

5 References

[1] "The Complexity of Cooperation: Agent-based Models of Competition and Collaboration", R. Axelrod, published by Princeton University Press, 1997, ISBN: 0691015678.

[2] "Increasing Returns and Path Dependence in the Economy", W. Brian Arthur, published by University of Michigan Press, 2000, ISBN: 0472064967.

[3] "Grid Computing: Making the Global Infrastructure a Reality", F. Berman, A. J. G. Hey, G. C. Fox, published by John Wiley and Sons, Ltd, 2003, ISBN: 0470853190.

[4] "Moore's Law", Intel WebSite, www.intel.com.

[5] "Chaos, Dynamics and Fractals, an Algorithmic Approach to Deterministic Chaos", J. L. McCauley, published by Cambridge University Press, 1993, ISBN: 0521416582

[6] "Is the Geormetry of Nature Fractal?" D. Avnir, O. Bihm, D. Lidar, O. Malcai published in Science, Vol. 279, No. 5347, Jan 1998.