

Ecological Analysis on Evolution of Information Systems

Ying Liu¹, Shu-ren Zhang², Mei-qi Fang¹

1 Information School, Renmin University of China
59# Zhong Guan Cun Street, Hai Dian District, Beijing, 100872, China
{ying.liu, fangmq}@ruc.edu.cn,

WWW home page: <http://ecolab.ruc.edu.cn>

2 School of Management, Hangzhou Dianzi University,
Hangzhou Zhejiang 310018, China
zhangshuren@gmail.com

Abstract. Information systems nowadays are growing and changing with a high speed. The flourish of system applications and participators imply opportunities and enthusiasm. Knowledge of the systems' evolution will be important to grasp the trend and so find a direction in the new era. To study the complicated and fast changing web environment, the article refers to ecological analysis. Two evolving models, agent emergence model and system integrated model, are studied. The evolving pattern of typical web applications is studied. According to the analysis, some potential opportunities are declared. Then collaborative structure of web components in integrated systems is analyzed based on network perspective. Popularity of web components and cluster structure of integrated systems, as well as their meanings, are discussed.

1 Introduction

Since Internet widely applied around the world, information systems connect and collaborate with each other. Internet links not only the static web pages, but different Information Systems (IS) as well. Recently great changes are taking place in the area of web. From a macroscopic perspective, some basic evolving rules exist in these changes. So there is a challenge for researchers to understand the existing tangled relationship on web and to find out the basic mechanisms. Once we grasp these rules, we can get prepared consciously before the changes take place in reality.

This article studies the evolution of information systems especially in recent fast changing Web2.0 era. [7] The approach of Information System Ecology is proposed. Some useful concepts from biology are referred. Evolution of information systems is

studied by analogy to that of species in ecology. Approach of network analysis is also referred. Two kinds of evolving models are discussed. For each model, supporting evidences in reality and potential applicable fields are presented.

2 Position of Information Systems Ecology

Ecological analysis on information systems includes the study of patterns forming, evolving, and also some interactive behaviors among systems. It is important to notice that this study is different from information ecology or organizational ecology. Comparison among the three fields may make the position of IS Ecology clear. Information ecology focuses on imitation, creation and diffusion of information from organizational perspective, and is mainly on the ecological behavior within an information system or virtual community. [1, 2] Organizational ecology utilizes biological analogy and statistical analysis, trying to understand conditions under which organizations emerge, grow, and die. [3, 4] So the basic unit in these two fields is organization of users. In contrast, basic unit here is information system. User communities now are treated as component of the information system or external environment, although they still play an important role in interactive collaboration and evolution of information systems.

These related research fields is helpful despite of differences. Their relationship can be shown clearly by the following Fig. 1. IS Ecology locates in the intersection of Information System, Organizational Evolution Theory and Ecology. We can draw on ideas from these related fields, especially, from study on organizational ecology, which covers organizational evolution theory and ecology and also already exhibits effectiveness. [5]

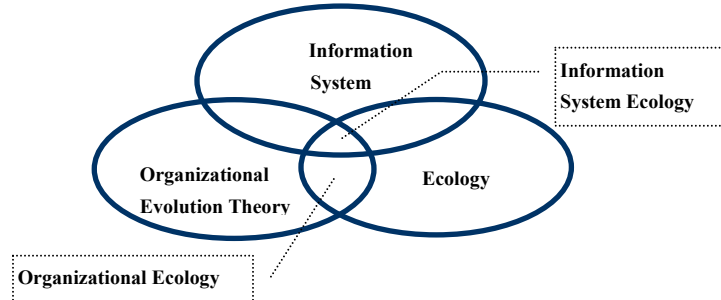


Fig. 1 Relation among three areas and the position of IS Ecology

3 “Cropper Principle” and Agent Emergence Pattern

Web 2.0 is mentioned more and more frequently as the name of a kind of system style. [7-9] Various web applications are flourishing nowadays and “grass root” plays a more important role on the stage of web. In this paragraph, the development of some typical technologies in Web2.0 will be studied. Some concepts and theories from biological fields enlightened our understanding of new phenomena in this area.

In our opinion, emergence of web applications recently can be considered analogical to the Cambrian Explosion to some extent. There are different explanations for the Explosion in biology. Steven M. Stanley of Johns Hopkins University provides a popular ecological theory. [10] He argued “cropping principle” provides a biological control. In one word, the theory explains that the organic diversity comes from “cropper” emergence. It means a “cropper” that feed on some existing species will emerge, along with the growing of species. And this process will continue, as the population of “cropper” grows to some extent, higher level “cropper” comes into being.

We find this theory can be also referred to explain the evolution of information systems. As we all know, web applications exhibit high diversity and extensive collaborating connections. While at the very beginning, the interaction among systems is limited when few of them adopted key technologies such as Ping, Trackback, RSS, Social Tags and etc. [11-13] We define these evolving pattern as “cropper principle” when a new system derives from another system that has developed to certain scale.

3.1 Agent Emergence Pattern of Ping

Take the evolution of Ping as an example, this process can be explained more clearly. (See Fig2) Ping is a kind of technology employed to detect real-time updating of social web sites, say group members' Blogs here. [14] At the beginning, a few intra-organization Blogs were inspected by a closed, directional Ping monitor. The Blog systems named as Ping source (now it refers to all systems that can send Ping messages) send Ping messages to Ping monitor. The monitor's interface opens to certain sources selectively, and system settings will have to be changed when changing cooperating group. So the cooperation among systems is local, closed and exclusive. (Fig. 2a) As the massive application of Blog systems, some Ping monitors begin to open the interface and allow Blogs add Ping connection by itself to inspect more systems distributed all around the world. So the open Ping monitor comes into force. We call this kind of open monitor as Ping pool. (Fig. 2b) The precondition for Ping pool is existence of large numbers of Ping sources. Weblogs.com is the pioneer of Ping pool. As it can conveniently collect a great deal of changing messages in time, other Ping monitors begin to imitate it. So a lot of Ping pool systems come into being. (Fig. 2c) The cooperation among systems evolves from N:1 to N:M pattern. Along with the similar increase in number of Ping pools, Ping sources that have to send message to multiple Ping pools also increase. It brings inconvenience to both sides, especially to the Ping sources. Those Ping messages sent out to multi-systems waste the bandwidth and reduce efficiency. So some Ping Pools change their strategy. They take charge of transmitting messages to other Ping Pools to attract more Ping sources' direct connections. We call this kind of Ping Pool as Ping Transmit Agent (such as PingOMatic, PinGoat etc.) (Fig 2d) Basic phases in the evolving process can be shown as following Fig. 2.

The emerging agent releases Ping source systems from low efficiency, as they needn't issue messages to many systems simultaneously. The same effect can be completed by only sending messages to a Ping Transmit Agent. The emergence of

agent changes system cooperating patterns from N:M to N:1:M. We can summarize this evolution pattern as agent emergence represented by “N:1 → N:M → N:1:M”.

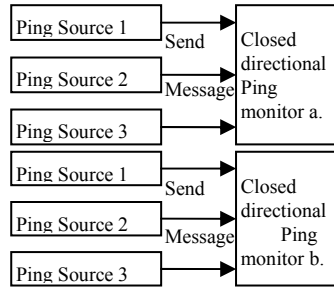


Fig. 2a Closed directional Ping Monitor

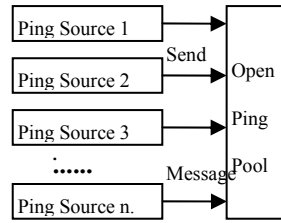


Fig. 2b Ping Pool

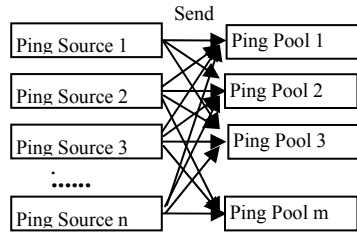


Fig. 2c Flourish of Ping Pool N:M

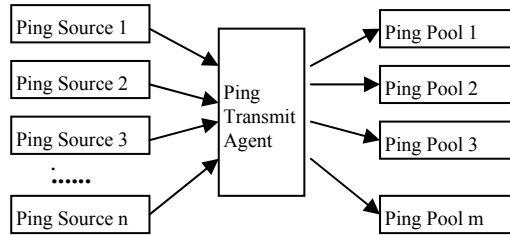


Fig. 2d Ping Transmit Agent N:1:M

Fig. 2 Evolution Process of Ping

3.2 Application Proposal

We can get some insight on system development from the evolving pattern study. Potential opportunities may be detected, which will help us get prepared consciously. In other words we can forecast what will probably emerge in web technology development by observing what already happened in related areas.

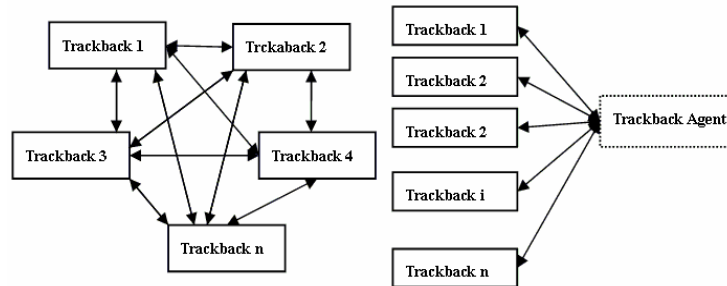


Fig. 3 Niche vanity in Track Back

For example, we focus on the technology of Trackback. Each system supporting Trackback develops interaction with other congeneric systems, which forms

bidirectional equal interaction among systems.(See Fig. 3 left) The interaction now is “many to many” pattern. According to “ $N:1 \rightarrow N:M \rightarrow N:1:M$ ” agent emergence pattern, Trackback agents will emerge sooner or later.(Fig. 3 right) Features and functions of the new system can be also predicted. It should be the agent systems that take charge of managing the Trackback interactions between two systems. Assisted by these agents requirement from a system is inducted to proper branch interface, and then sent to the objective system through established connection. Trackback agent will be useful in centralized verification, like cooperating anti-spam, reputation assessment and so on.

4 Network Analysis and System Integrated Model

Open fabric is the key to gain adaptability and also the key to survive for information systems nowadays. Special web components are designed to be integrated easily. Some components including Google API, Yahoo API, and etc. are widely applied in many systems. The principle is to integrate functional components in a system in stead of redesigning from nothing. In this way, new integrated systems emerge quickly. We define this phenomenon as system integrated model.

The collaborative relationship of the components can be studied by network analysis approach. Network analysis is a kind of structure analysis. Graph theory provides necessary theoretical technology. [15] It provides us a brand new perspective to understand the environment, and it also introduces a set of new concepts, methods and tools. In this perspective, the key to understand behavior is the holistic structure, the opportunities and restricts from position. To study the collaboration of components, here nodes represent components and a tie occur when there exists a system combined the two components represented by the nodes on both ends. Then we get the collaborative network. So the topology can be visualized, and some quantitative properties like centrality and cluster structure can be calculated and analyzed. Evolving trend of the network also can be studied based on time related data. The analysis may help us making wise decision when choosing components for a new system.

An on line system — Programmable Web —collects mainstream programmable web components and mixed applications. (It can be accessed at <http://www.ProgrammableWeb.com/mashup>.) We collect the real data from the webpage to build up the network. The sampling was conducted in a series of three times. The analysis employed network analysis tool – UCINET. [16] We’ll show our study on centrality, sub-group cluster of the component collaborative network.

4.1 Centrality

From network analysis perspective, centrality is an important quantity indicator to measure the power of a node from its position. There are many kinds of centrality in different level, for example Degree Centrality, Betweenness Centrality, Closeness Centrality, and Harmonic Closeness Centrality.

As our interest is mainly on the evolving pattern of the whole network, we refer to the Harmonic Closeness. We calculate the closeness centrality value in the way

Stephenson and Zelen proposed, as it provides the normalized centrality information for each node in the whole network context. Centrality of a vertex i is the harmonic mean of the measures between i and all other vertices in the network. [17] We list the rank of this value for some popular web components in Table 1. The number in the table cell means the rank value of the harmonic closeness centrality calculated for the component, and “1” means the highest in the network. The first column is the name of the component, and the column 2, 3, 5 correspond to the data collected in three different dates, which is shown in the first row. The column 4 and 6 display the different in rank number in the latest two investigation samples.

Table 1. Harmonic closeness Rank

	Dec 5 th	Jan 15 th	Difference	Feb 1 st	Difference
Amazon	1	2	-1	1	1
Flickr	2	1	1	2	-1
GoogleMap	3	3		3	
Technorati	7	4	3	4	
del.icio.us	4	5	-1	5	
eBay	14	6	8	6	
GoogleAPI	6	7	-1	7	
A9	27	8	19	8	
VirtualEarth	5	9	-4	9	
YahooMaps	23	17	6	10	7
FeedMap	8	15	-7	12	3

From Table 1 we can see the changing position of the components. We highlight the obvious difference. A9, Yahoo Maps, eBay and Technorati are the components growing fast, and FeedMap and VirtualEarth are the two descend greatly in the list. From the changing value of the harmonic closeness of those web components, we can imagine corresponding change in their popularity and influence. The components that are going up in the rank list represent the trend of technology to some extent. It displays the choices of most integrated systems. Based on this observation and analysis, we can select a component more purposefully when there are some options.

4.2 Cluster Structure

Based on the study of web components, we can further analyze integrated information systems. Cluster study of the systems will help us detect the “species”. Fig. 4 is a visualization of systems cluster study. Nodes in the figure are integrated systems building on web components discussed above. The UCINET algorithm “Cliques” produces a census of all cliques, which means maximal complete sub-graph. Frames mark in Fig. 4 is according to those results. “Species” are defined in this way here. This visualization makes the structure obvious and easy to understand. It is helpful for information systems to find their latent rivals or cooperators. In a more complicated and confusing business environment, this ability is more and more important for practitioners of e-commerce.

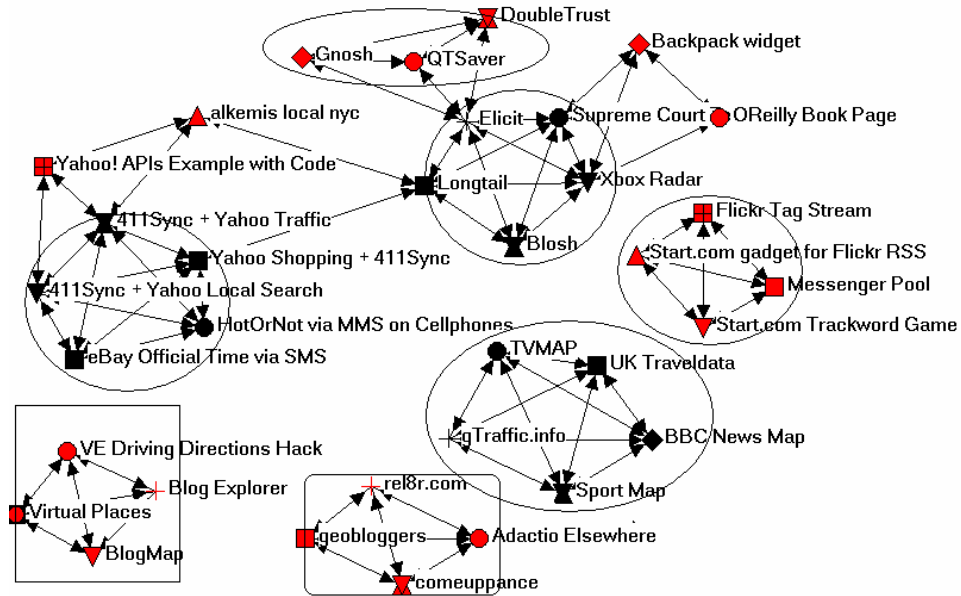


Fig. 4 System ecology and species

5 Conclusions and Further Studies

The fast developing applications of information systems attract our interests. Key technologies related applications and integrated systems are studied. By ecological analysis, we found “cropper principle” works in the evolution of Ping related systems. Based on this analysis, the potential opportunities in Trackback related systems are discussed. Further, according to the trend of open fabric in information systems, our study focuses on integrated systems. Network analysis approach is referred to study the collaboration of web components and also the structure of integrated systems. The popularity of different component at different time is studied through harmonic centrality. And the whole structure of integrated systems is studied by cluster analysis.

The ecological analysis on information systems helps us to grasp the rule of their evolution. To some extent it provides us capability to forecast the changing. The high speed developing and changing of information systems is still on going, the capability can instruct our direction.

Further studies in this area need more of our work. There are lots of useful concepts and theories from ecology can be introduced to information systems analysis. Systemic investigation could be interesting. And the new systems is growing and changing quickly, so we have a growing dataset. We should develop some automatic data-gathering technique to collect useful information from the on-

line resources. Some adaptation of network analysis is also expected to fit for the certain purpose.

References

1. T.H. Davenport, L. Prusak, *Information Ecology: Mastering the Information and Knowledge Environment*, Oxford University Press, New Your, NY, (1997)
2. Y. Malhotra, *Information Ecology and Knowledge Management: Toward Knowledge Ecology for Hyperturbulent Organizational Environments, Encyclopedia of Life Support Systems (EOLSS)*, UNESCO/EOLSS Publishers, Oxford, UK. (2002)
3. M.T. Hannan, J. Freeman, *Organizational Ecology*, Cambridge, MA: Harvard University Press. (1989)
4. G.R. Carroll, M.T. Hannan, *The Demography of Corporations and Industries*, Princeton, NJ: Princeton University Press. (2000)
5. J.A.C. Baum, J.V. Singh, Organizational Niche and the Dynamics of Organizational Founding. *Organization Science*, 5(4):11. pp 483-501 (1994)
6. E.R. Pianka, Competition and niche theory, *Theoretical Ecology*, Second Edition edited by R. M. May, Chapter 8 (Blackwell, 1981) pp. 167-196
7. T. O'Reilly, What Is Web 2.0, Design Patterns and Business Models for the Next Generation of Software (2005, 9),
<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
8. T. O'Reilly, Web 2.0 Compact Definition, (2006, 10)
http://radar.oreilly.com/archives/2006/12/web_20_compact.html
9. D. Hinchcliffe, The State of Web 2.0, 2006, 4.
http://web2.wsj2.com/the_state_of_web_20.htm
10. S.M. Stanley, An Ecological Theory for the Sudden Origin of Multicellular Life in the Late Precambrian, PNAS 1973 70: 1486-1489 *Proceedings of the National Academy of Sciences* (1973)
11. "TrackBack Technical Specification",
(<http://www.movabletype.org/docs/mtrackback.html>)
12. T. Hammond, T. Hannay, and B. Lund, The Role of RSS in Science Publishing, "Syndication and Annotation on the Web", *D-Lib Magazine*, Nature Publishing Group, (2004, 12) (<http://www.sythe.org/rss/12hammond.html>)
13. A. Mathes. Folksonomies - Cooperative Classification and Communication Through Shared Metadata [EB/OL], <http://www.adammathes.com/academic/computer-mediatedcommunication/folksonomies.html>. 2004)
14. M. Jensen, E. Alternatives, A Brief History of Weblogs, *Columbia Journalism Review*, 2003 (5)
15. S. Wasserman, K Faust, D. Iacobucci, *Social Network Analysis: Methods and Applications*. Cambridge, Cambridge University Press (1994)
16. S.P. Borgatti, M.G. Everett, and L.C. Freeman, *Ucinet 6 for Windows*, Harvard: Analytic Technologies. (2002)
17. R.A. Hanneman, and M. Riddle, *Introduction to social network methods*, Riverside, CA: University of California, Riverside (2005)