Structure and Collaborative Aspects of Internet Standards

Mehmet Gençer
Istanbul Bilgi University,Department of Computer Science
Kurtuluş Deresi Cad. No:47 Dolapdere, Istanbul, Turkey
mgencer@cs.bilgi.edu.tr

Abstract

Standardization efforts have become an increasingly central issue as business interoperability demands mount. However research on the development process of standards is limited. This study investigates the organizational features of Internet standards developed and published under coordination of the Internet Engineering Task Force. In addition to longitudinal changes in standards development process, patterns of links between standards and authors, and extend of deliberation and collaboration were investigated.

1. Introduction

Development of computer communication standards dates back to late 1960s, motivated by data portability and digital communication needs. With formation of The Internet Engineering Task Force(IETF) in 1986, it became the steering body for development of Internet related interoperability standards. With early inclusion of non-governmental contributors, IETF managed to stand to its intended position in provision of Internet standards. Today the corpus of standards produced by IETF contains over four thousand 'request for comment' documents(RFCs), and it remains as one of the central bodies alongside with Institute of Electrical and Electronics Engineers(IEEE), International Standards Organization(ISO), and others, for development of standards aimed at provision of interoperability in the larger information technology domain.

The structure, procedures and objectives of standardization efforts are explicit and well established[1]. While respecting openness and aiming wide collaboration, standards must be coherent, well tested, but they also must arrive timely to facilitate institutional convergence. These requirements pose high demands in efficiency of collaboration in producing standards.

General requirements for standardization efforts has been considered in research[6], as well as more general aspects of IT business such as increased influence of collaborative knowledge creation [10], alliances[9, 4], and consequences of externalities[5]. However, empirical research and case studies specifically targeting standards development and network effects on the development process received attention only recently [7, 3, 13, 11, 8].

On the other hand social networks and alliances has been

studied extensively in social sciences [12, 9, 2], and methods were developed to qualify structural features of network relations. Drawing on these approaches, this study first attempts an extensive quantitative analysis of the corpus of IETF standards to illuminate commonalities and their longitudinal change, and then focuses on understanding emergent patterns in relations between standards and their authors to their network positions, in relation to possible consequences for standards development planning. While traditional statistical tools are sufficient for the first part, methods from social network analysis(SNA) are employed for the second part of investigation.

The next section presents a description of IETF standards and data used in the study, followed by a presentation of research methodology including SNA methods deployed in the analysis. Results of statistical analysis and network analysis are presented in the subsequent sections, followed by a summary of conclusions and discussion of consequences for standards development planning.

2. IETF standards and RFC corpus

Since its foundation in 1986, IETF has been the primary initiative for many Internet related standards, including those for the Internet protocol and other TCP/IP stack protocols, e-mail exchange formats, and many more. Some historic protocols were also formalized and made available[1]. IETF standards proceeds through publication of 'request for comment'(RFC) documents for wide circulation. Through inclusion of comments and implementation results from relevant parties, the process is iterated through publication of updates until necessary level of adoption is signaled by the community(see [1] for a detailed description of IETF procedures).

This iterative process is reflected in the classification of RFCs. Some RFCs are tagged as *Experimental* or *Informational*. Another group, *Best Current Practice* RFCs "standardize the results of community deliberations about statements of principle or conclusions about what is the best way to perform some operations"[1]. The final group, *standard* RFCs, are the specifications adopted as Internet standards. Number of RFCs in different classes and sub-classes are shown in Table 1; the order suggests an increasing level of class importance.

Class	**Number of RFCs**
UNKNOWN	909
HISTORIC	200
EXPERIMENTAL	264
INFORMATIONAL	1430
PROPOSED STANDARD	1454
BEST CURRENT PRACTICE	131
DRAFT STANDARD	122
STANDARD	87

Table 1. Number RFCs in different classes

Many RFCs cross reference others. These relations must be considered as they indicate influence of RFCs referred to, as in citations in the scientific literature. In addition many authors are involved in development of more than one RFC. These co-authoring ties capture the structure between the authors. Nevertheless these relations provide a good basis to investigate structures within the IETF standards development.

[1] All RFCs are accessible via IETF website at http://www.ietf.org

3. Research method and social network analysis

Development of Internet standards is a process informed by wider developments in information technology, and in-depth understanding of this process requires intensive case studies. However, in this study we attempt a rather extensive and descriptive investigation of standards development. For this purpose output volume of standards development efforts, team sizes of development efforts, and their change over time must be considered to understand interaction of socialization level and development efficiency.

In addition links between RFCs(citation links) and authors(co-authoring links), and their longitudinal change can provide information on effects of socialization. For this purpose an essential SNA indicator, in-degree prestige, can be employed. In-degree prestige is simply the number of links to a node and higher values indicate higher influence of a node in our network. In-degree prestige of a node, i, in a network with N nodes, can be formulated as follows[12]:

$$P_i = \sum_{k=0}^{N} l(k,i)$$

where $l(k,i)$ is defined for RFC citation network as follows:

$$l(k,i) = \begin{cases} 1 & \text{if RFC}k \text{ refers to RFC}i \\ 0 & \text{otherwise} \end{cases}$$

and similarly for co-authoring network:

$$l(k,i) = \begin{cases} 1 & \text{if } k \text{ has authored any RFCs with } i \\ 0 & \text{otherwise} \end{cases}$$

One other measure regarding relations used in our study is the graph density, defined as the ratio of possible lines to which are actually present[12]. Since citation relations between RFCs are directed relations, while co-authoring relations between authors are not directed, we will use different formulations for each. For directed RFC reference relations, the graph density is computed as follows:

$$\Delta = \frac{L}{N(N-1)}$$

where L is the number of relations present, and N is the number of RFCs. Similarly the density of undirected graph showing co-authoring relations is computed as follows:

$$\Delta = \frac{L}{N(N-1)/2}$$

In addition to metrics and statistics mentioned above, graph visualization tools are helpful for contextual understanding of networks. One such tool, Pajek[2], is used for visualization purposes in our study.

4. Research Results

Variety of analysis presented below can be laid out in two categorizations. First is that some findings are related to RFCs themselves and their citation relations, whereas others to the authors and co-authoring relations. Second, where some measures consider to-date snapshots of the standards corpus, others relate to changes in the corpus through time. Findings are presented below to provide a comparative view of standards and authors.

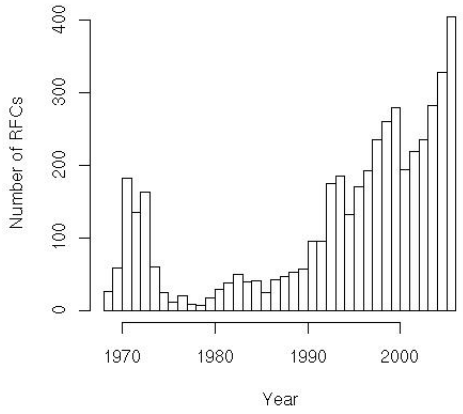
4.1. Output volume and collaborative nature of standardization efforts

Figure 1(a) shows number all RFCs appeared every year. A steady increase trend following establishment of IETF in 1986 is apparent in the total number of RFCs published every year. However we do not see such a trend in the case of *standard* class RFCs. Thus we can say that required level

[2]Pajek program and documentation are available at http://vlado.fmf.uni-lj.si/pub/networks/pajek/

of deliberation is increasing in the standardization process. In other words, more and more *informational*, *best current practice*, etc., class RFCs are needed for reaching consensus on a *standard* class RFC.

Figure 1(b) shows the number of authors of RFCs changing through years. Here, too, we see an upwards trend. Therefore we can say that not only the level of deliberation, but also the level of collaboration is increased.



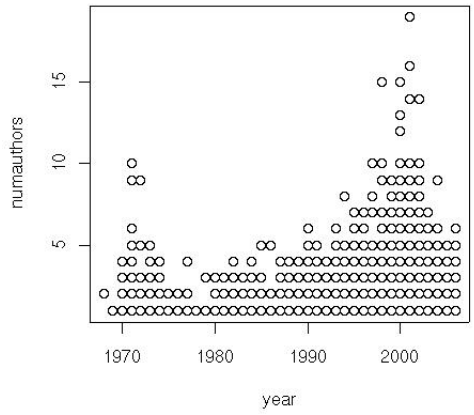


Figure 1. Number of standards and authors.

4.2. Effects of team size and author relations

Figure 2 shows number of authors for different classes of RFC. *Standard* class RFCs appear to have lower number of authors compared to others. This is possibly due to the fact that level of consistency required for this class renders higher number of authors un-workable.

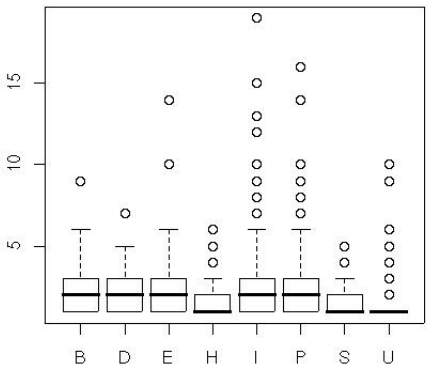


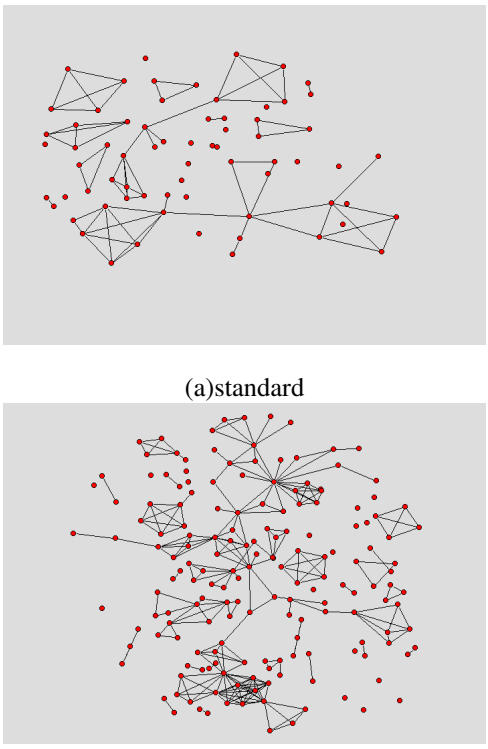
Figure 2. Number of authors for RFC types.

Effects of socialization can be hinted upon by looking at the number of ties between authors and their influence in the standardization process. Figures for number of co-authors worked together(in-degree) and number of RFCs authored for authors of *standard* and *best current practice* class RFCs are analysed. In the case of *standard*s, we observe a considerable relation between the two. The correlation of two variables is 0.726 for this class. On the other hand *best current practice* class has a lower correlation level of 0.554. Although these results are not sufficient alone for deeper conclusions, we may note that higher levels of social ties to other authors goes hand-in-hand with the level of influence in standards development, and this is more apparent in

the case of standards, whereas other class RFCs appear to be more heterogeneous in this respect. Therefore involvement in RFC publications appears to be more related to social ties in the case of higher(*standard*) class, compared to other(lower) classes. Although no causal relation is suggested by these results, possibility of a dual positive feedback needs to be further explored.

4.3. Network structure of author relations

Figure 3(a) and (b) shows a visualization of co-authoring relations for authors of *standard* and *best current practice* class RFCs. The graph densities are 0.087 and 0.037, respectively. The difference in densities is not surprising due to different number of nodes in two graphs. In both cases, one observes certain 'islands' of activity, and certain authors who facilitate relations between them. Therefore there is no reason to think that collaboration patterns are different for two cases.



(a)standard

(b) best current practice

Figure 3. Co-authoring relations between authors of RFCs.

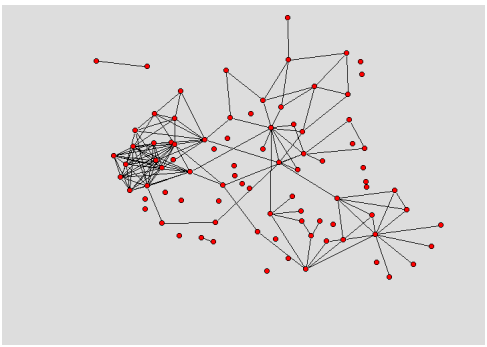
4.4. Citation relations between RFCs

Figure 4(a) and (b) shows a visualization of referral links between RFCs for *standard* and *best current practice* class RFCs. The graph densities are 0,022052927 and 0,014445097, respectively. Once again the difference is possibly due to difference of the graph size. However the structure of relation graphs for *standard* and *best current practice* classes appear to have differences. In the case of *standard* class, apart from one region where relations are concentrated[3], one sees many work domains that are occasionally tied to each other. However in the case of *best current practice* class, few RFCs have a high number of ties. Distribution of in-degrees for these cases confirm this finding. One possible reason for this concentration in the case of *best current practice* class is the fact that this group spans a smaller time interval, and hence number of conjectural foci of study are smaller. However a more comprehensive explanation begs for further study of these cases.

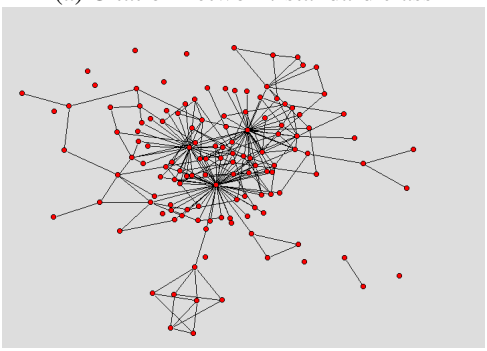
[3]This region was identified elsewhere[8] as standards related to network management protocols

5. Conclusion

Our investigation of Internet related RFC publications of the IETF illuminates several facts about the standards development process. First is that the proportion of non-*standard* class publication activity is increasing. While there is no upwards trend in number of *standard* class RFCs, the over-



(a) Citation network: standard class



(b) Citation network: Best current practice class

Figure 4. Referral relations between RFCs.

all number of RFCs published every year has steadily increased. This is an indicatior that level of deliberation required for reaching a consensus on standards is increasing.

Second, the extend of community collaboration is increased. This is pointed by the increase in the number of authors involved in publication of RFCs. However it is worth noting that extend of collaboration is lower for *standard* class. One possible interpretation is the fact that consistency required for this class is too high to accomodate large number of influencing authors.

Third, a juxtaposition of the authors co-authoring relations and number of RFCs they are involved, we have found a considerably high level of correlation between the two, and more so in the case of *standard* class. However we make no claims about causal relationship between the two. This aspect of the process requires in-depth, qualitative case studies. We have also compared the patterns of the co-authoring relations for different classes, but have found no significant differences between them.

Finally, we have looked at the patterns of citation relations between RFCs. Here we have found significant differences between *standard* and *best current practice* classes. Very few RFCs are found to have outstandingly central position in the latter. The interpretation we offer is the lower number of attention foci due to lower time interval coverage of data in this case.

We believe further studies concerning developmental patterns in similar corpuses would be valuable for planning and monitoring purposes. Standards development requires a fine balance between demands of timely delivery, wider collaboration, and excellent technical consistency. The trends highlighted in this study point to hardening of certain conditions. A careful consideration of relevant findings for planning and policy making in standardization efforts will be useful for achieving timely delivery of standards.

References

[1] S. Bradner. *The Internet Standards Process (RFC2026)*, 1996.

[2] R. S. Burt. *Structural Holes*. Harvard University Press, Cambridge, 1992.

[3] B. Choi, T. S. Raghu, and A. Vinze. An empirical study of standards development for e-businesses. In *Proceedings of the 39th Hawaii International Conference on System Sciences*, volume 6, page 139a, Los Alamitos, CA, USA, 2006. IEEE Computer Society.

[4] P. A. David and L. C. Keely. The endogenous formation of sciencetific research coallitions. *Economics of Innovation and New Technologies*, 12:93–116, 2003.

[5] B. Demil and X. Lecocq. Neither market nor hierarchy nor network: The emergence of bazaar governance. *Organization Science*, 27:1447–1466, 2006.

[6] D. L. Garcia. Standard setting in the united states. *Journal of the American Society for Information Science*, 43:531–537, 1992.

[7] R. Garud, S. Jain, and A. Kumaraswamy. Institutional entrepreneurship in the sponsorship of common technological standards. *Academy of Management Journal*, 45:196–214, 2002.

[8] M. Gencer, B. Ozel, V. S. Tunalioglu, and B. Oba. *Open Source Systems*, chapter Organization of Internet standards. Springer, Boston, 2006.

[9] R. Gulati, N. Nohria, and A. Zaheer. Strategic networks. *Strategic Management Journal*, 21:203–215, 2000.

[10] G. K. Lee and R. E. Cole. From a firm-based to a community-based model of knowledge creation. *Organization Science*, 14:633–649, 2003.

[11] J. V. Nickerson and M. zur Muehlen. Defending the spirit of the web: Conflicts in the internet standards process. In J. L. King and K. Lyytinen, editors, *Proceedings of the Workshop on Standard Making*, pages 56–69, 2003.

[12] S. Wasserman and K. Faust. *Social network analysis: Methods and applications*. Cambridge University Press, Cambridge, 1994.

[13] M. Xia, K. Zhao, and M. J. Shaw. Open e-business standard development and adoption. In J. L. King and K. Lyytinen, editors, *Proceedings of the Workshop on Standard Making*, pages 56–69, 2003.