

## **Economic Modeling of Social Network Services**

Hyunwoo Park

hwpark@ischool.berkeley.edu

School of Information

University of California, Berkeley

### **Abstract**

In this paper, social network services are economically modeled. Many studies have attempted to describe the nature of social network structure, but few studies quantifying the incentives for individuals to join a social network have been released. I argue that the combination of existing research on network topology and quantified model of incentives for joining a social network service can better describe micro interactions within a social network service. In addition to individual-level interaction modeling, initial user base of social network sites are modeled as well. The model of interactions describes dynamic part of social network services, and the model of initial user base is similar to boundary conditions. I argue that understanding these two at the same time will provide better understanding of how social network services prosper and perish.

**Keyword:** social network, economic modeling

## 1. Background

As the Internet emerges, many social network services have come out since the 90s. These services have gained millions of users during last decade.<sup>1</sup> Latest popular services are MySpace, Facebook and Twitter. These services generally allow users to maintain relationship through the Internet.

Many studies have analyzed structural characteristics of social networks and have proposed social implications. Wasserman and Faust<sup>2</sup> wrote a comprehensive handbook for social network analysis. It employed graph notation to describe and analyze social networks. In terms of social implications, Granovetter argued that weak ties play a crucial role of diffusing information across social networks.<sup>3</sup>

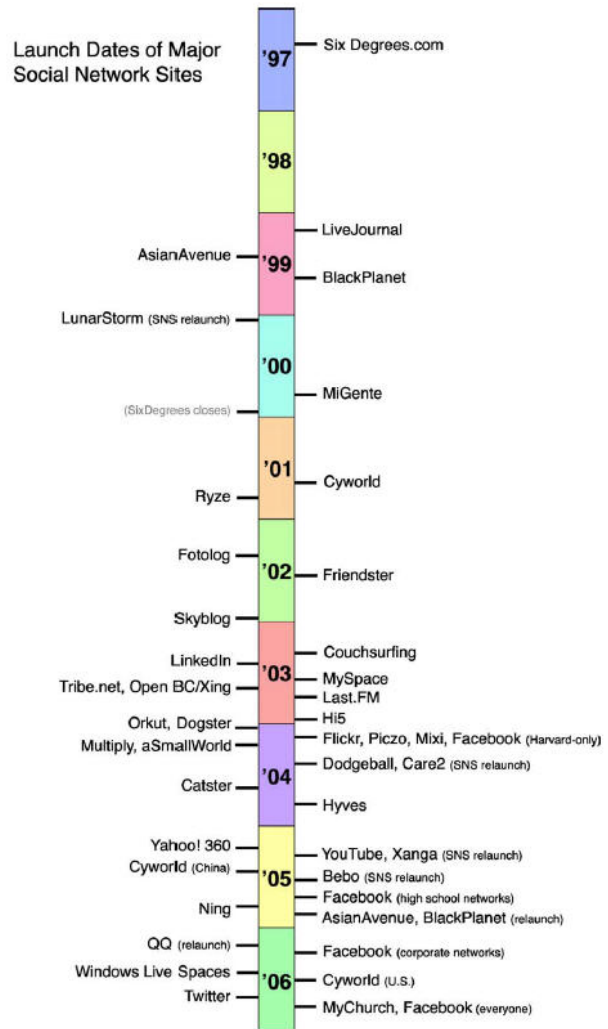


Figure 1. Timeline of social network sites

Some scholars have also proposed models of technology diffusion processes. Geroski summarized models of technology diffusion.<sup>4</sup> Epidemic model and probit model are mainly discussed in the paper. However, these models focus on macro level observation of diffusion processes.

On the other hand, little literature on modeling economic behavior within social networks from microscopic point of view has been released. In this paper, I suggest economic models of interactions within social networks and initial state of such services.

<sup>1</sup> Boyd, D. & Ellison, N., Social Network Sites: Definition, History, and Scholarship, *Journal of Computer-Mediated Communication* 13 (2008) 210-230

<sup>2</sup> Wasserman, S. & Faust, K., *Social Network Analysis: Methods and Applications*, New York, NY: Cambridge University Press.

<sup>3</sup> Granovetter, Mark. 1973. "The Strength of Weak Ties." *The American Journal of Sociology* 78:1360-1380.

<sup>4</sup> Geroski, P.A. 2000. "Models of technology diffusion." *Research Policy* 29:603-625.

## 2. Framework

In this paper, social network services are defined as web-based services offering a list of other users with whom they share a connection and an ability to traverse the list.<sup>5</sup> Such services do not usually enable relationships between strangers, rather they help users visualize and articulate existing offline relationships.<sup>6</sup> Therefore, existing offline relationships and connections between individual and social network services should be differentiated. I argue that social network services should be modeled as bipartite graph rather than homogeneous unipartite graph. The conceptual diagram is presented below.

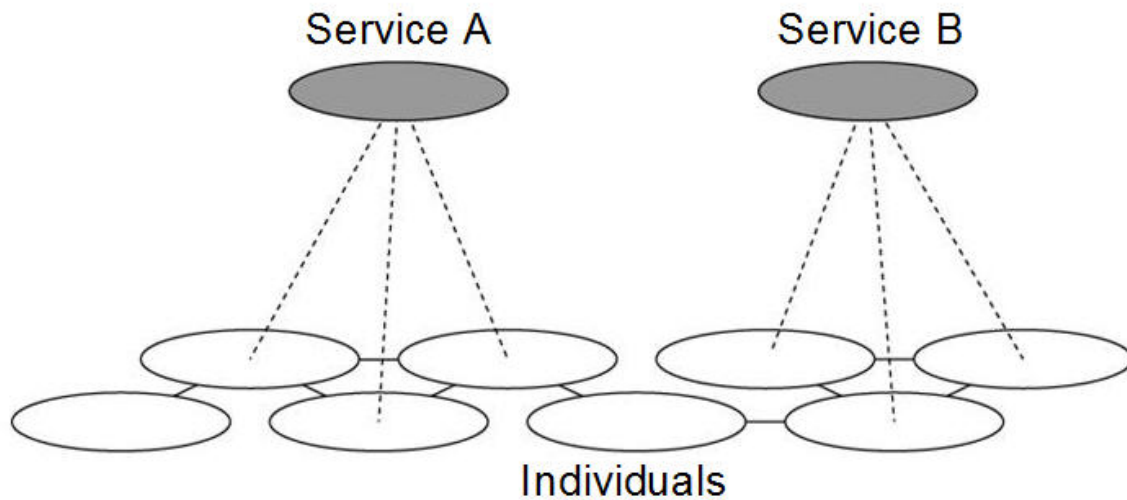


Figure 2. Framework of modeling social network services

In the diagram above, white circles represent individuals in a social network and gray circles conceptualize social network services. Vertices can be represented as follow.

$$V_{individual} = \{v_i \mid i = 1 \text{ to } \# \text{ of individuals}\}$$

$$V_{service} = \{v_i \mid i = 1 \text{ to } \# \text{ of services}\}$$

Meanwhile, the solid lines connecting white circles mean offline relationship which has existed even without social network services. The dotted lines are membership relations between individuals and services. The strength of offline relationship can differ, so weighted edge is appropriate to describe the

<sup>5</sup> Boyd & Ellison

<sup>6</sup> Haythornthwaite, C. (2005), Social networks and Internet connectivity effects, *Information, Communication, & Society*, 8(2), 125-147

relationship. However, membership can be described as an on-and-off notation. In sum, edges can be described as follow.

$$E_{individual-individual} = \{e_{ij} \mid v_i \in V_{individual} \wedge v_j \in V_{individual}, e_{ij} = 0 \dots \text{strength of relationship}\}$$

$$E_{service-service} = \{e_{ij} \mid v_i \in V_{individual} \wedge v_j \in V_{service}, e_{ij} = 0 \vee 1\}$$

The description above employed two different types of vertices and edges. The study of Eubank et al<sup>7</sup> used similar approach to describe a social network. They used two types of nodes: people and locations.

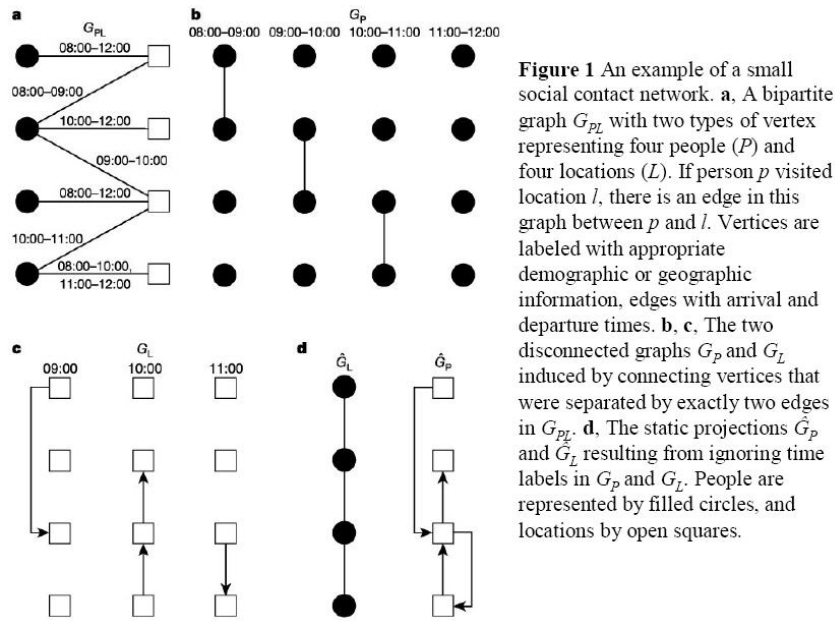


Figure 3. Example of using bipartite graph in social network analysis

In the context of modeling social network services, two parts should be considered. One is how individual decides which service to join. The other is how to describe the initial state of social network services.

<sup>7</sup> Eubank et al, Modeling Disease Outbreaks in Realistic Urban Social Networks, Nature 429, 180 (2004)

### 3. Modeling Interactions

There are two types of individual-level interactions in this modeling. One is the link between individuals, which is offline relationship that exists even without social network services. The other is the link between individual and service, which is membership.

#### *a. Individual-Individual Interaction*

Several researchers have identified and tried to mimic the nature of real world social networks. The simplest network topology is a random topology. Waxman<sup>8</sup> is one of the first researchers who investigated the characteristics of random networks. To generate Waxman topology, a set of vertices is first geographically placed. Then, the probability of having a link between two vertices depends on the distance between the two. Since the places of nodes are decided randomly, Waxman topology is one of random networks.

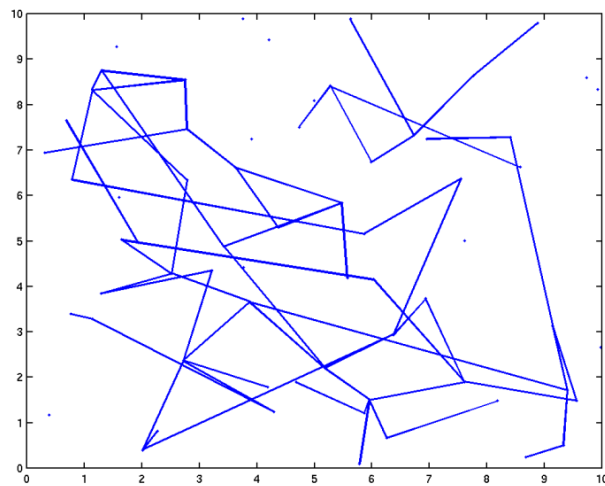


Figure 4. Waxman random network topology<sup>9</sup>

---

<sup>8</sup> Waxman, B. 1988. Routing of multipoint connections. Selected Areas in Communications. IEEE Journal on.

<sup>9</sup> <http://www.math.uu.se/research/telecom/software/stgraphs.html>

Although Waxman topology is one of the first attempts to systemically generate network topology, it is not appropriate for describing social networks because some nodes have much more connections than other nodes in the real world. As an alternative topology, Barabasi-Albert topology is proposed.

Barabasi-Albert topology<sup>10</sup> is generated using incremental growth model. New nodes are added to existing network one by one. When a new node is added to an existing network, the probability whether a new link is created between the new node and existing nodes depends on the number of existing links belonging to existing nodes. It is called preferential creation of edge principle. As a result, Barabasi-Albert topology is scale free network, which means that its degree distribution follows a power law.

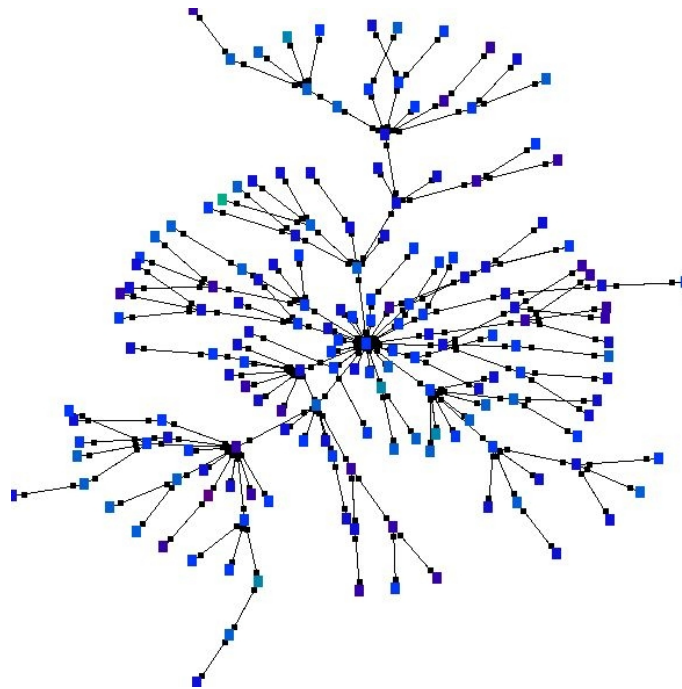


Figure 5. Example of Barabasi-Albert scale free network<sup>11</sup>

Barabasi-Albert network is one of small-world networks. There are several ways to generate small world topology. Another algorithm starts with  $N$  nodes from the beginning. Each node equally has  $K$  connections. With given probability  $p$ , whether to break each link is determined. If a link is broken, a new link is created with another node. The network generated in this way also exhibits small world features.<sup>12</sup>

---

<sup>10</sup> Barabasi, A., and Albert, R. 1999. Emergence of Scaling in Random Networks. Science.

<sup>11</sup> Hoffman, Jager, & Von Eije, Social Simulation of Stock Markets: Taking It to the Next Level, Journal of Artificial Societies and Social Simulation vol. 10, no. 2, 7

<sup>12</sup> Li D. & Du Y., Artificial intelligence with uncertainty, Chapman & Hall/CRC, p.333

There is another approach to generate networks by setting a game situation. Skyrms and Pemantle<sup>13</sup> suggested a dynamic model of social network formation. In this model, each agent play a game with other agent. Two types of games are given: friendly and hostile games. In a friendly game, one or both of agents get positive payoff per each interaction. Conversely, in a hostile game, one or both of agents get negative payoff. The game payoff tables are as follow.

Friends I	Host	Friends II	Host	Enemies I	Host	Enemies II	Host
Visitor	(1, 0)	Visitor	(1, 1)	Visitor	(-1, 0)	Visitor	(-1, -1)
<b>Friendly game</b>				<b>Hostile game</b>			

Figure 6. Game payoff tables summary

*b. Individual-Service Interaction*

While there are many studies about individual-individual interaction, few studies have investigated the relationship between individual and social network service. The interaction between individual and service can be economically modeled from two perspectives: utility and cost. Utility is positive incentive for individuals to join a network, and cost is negative incentive.

*i. Utility function*

Utility perspective describes how individual obtains benefit from joining a social network service. The first obvious factor affecting utility of individual is the number of friends who already joined the network. As the number of friends in the same service gets larger, an individual will have greater utility by joining the network. Not only does the number of direct



Figure 7. Japanese blog service, Ameba

<sup>13</sup> Skyrms B. & Pemantle R., A dynamic model of social network formation, PNAS vol. 97, no. 16, 2000

friends affect the utility, but the number of friends' friends can have an effect. However, it is reasonable to assume that the latter has less impact on the utility than the former.

I would suggest another factor, the number of celebrity in the network. For example, one of Japanese blog service provider called "ameba" differentiates itself by hosting blogs of celebrities. (Figure 7) Celebrity can be defined as a person who is known to many people but knows relatively little number of people. Strength of celebrity can also be defined as the number of people who know the person.

In sum, if the decrease of impact as depth of friends becomes deeper is assumed as an exponential function, utility function of agent  $i$  can be defined as follow.

$$u_i = p \sum_j x_j e^{-aj} + q \sum_k y_k$$

$$\left( \begin{array}{l} a, p, q: \text{constants} \\ x_j: \text{\# of friends in depth } j \\ y_k: \text{strength of celebrity } k \\ j: \text{\# of depth of friends of individual } i \\ k: \text{\# of celebrity in the network} \end{array} \right)$$

ii. Cost function

On the other hand, cost side impedes individuals to hesitate to join a service. Individuals usually have to pay attention, time and effort, or even real money as joining a social network service. Service providers can convert users' attention to monetary value by placing an advertisement. Service lifecycle of social network sites can be generally divided as three stages: joining, using, and quitting.

Joining cost is one-time cost which a user should pay when starting to use a service. Checking end-user license agreement or privacy policy is one of examples of joining costs. Setting up login id, password and profile can also be considered a part of joining cost. Therefore, I would argue that joining cost is a fixed cost rather than a variable cost. Joining cost differs from service to service. If a service provider offers an easier way to join, its joining cost should be regarded lower than other service providers' offering.

Using cost is a cost which occurs as user uses a service. Effort of getting notified from friends' status change or traversing one's friend network can be regarded as a using cost. Since using cost is charged as service is used, the nature of cost tends to be variable than fixed. One more thing to consider for the using cost is learning curve. As people become familiar with a service, the using cost should be lowered. Learning curves generally follows exponentially decreasing trend.<sup>14</sup>

Last, I argue that exit cost in a social network service is virtually zero. One can think that losing the network data established in a service should be counted as exit cost. However, the benefit from establishing a social network in a service is included in the utility function part previously. If it is counted as an exit cost as well, it is double counted.

In addition, website design and usability affects all three costs. If it is difficult for users to find what they want to do, all tasks from joining to quitting become harder. Therefore, it is reasonable to add usability factor to the total cost function. As a result, the cost function can be described as follow.

$$c_i = b_1 c_{join} + b_2 c_{use} e^{-r_i t} + b_3 c_{design}$$

$$\left( \begin{array}{l} b_1, b_2, b_3 : \text{constants} \\ r_i : \text{learning curve coefficient for individual } i \\ c_{join} : \text{joining cost} \\ c_{use} : \text{using cost} \\ c_{design} : \text{design \& usability cost} \end{array} \right)$$

Combining utility and cost functions, I propose a general formula for describing net value which individual acquires by joining a social network service.

$$v_i = u_i + c_i = \left( a \sum_j x_j e^{-bj} + c \sum_k y_k \right) + \left( b_1 c_{join} + b_2 c_{use} e^{-r_i t} + b_3 c_{design} \right)$$

---

<sup>14</sup> Yelle, L. E., The Learning Curve: Historical Review and Comprehensive Survey, Decision Science, 1979

## 4. Modeling Initial State

To fully describe social network services, not only individual-level modeling but also initial state modeling should be built. Such a factor which will be described here is the characteristics of initial member population. In social networks, the network effect and positive feedback mechanism rule. The more population a service has, the greater attraction it obtains. However, no service can start with zero population. Some early adaptors should bear negative net value at the beginning, so that the service takes off. Therefore, initial population should be considered a given factor in order to model social network services. The number of members at the initial stage and the distribution style are to be further discussed.

### *a. Size of initial user base*

The number of members at the initial stage can be represented either as absolute number of members or as percentage to total population. I argue that percentage notation is more appropriate to describe the size of initial population, because that approach is more scalable as the size of given offline network changes. Therefore, the size of initial user base can be states as follow.

$$S_i = \frac{n}{N}$$

$\left( \begin{array}{l} S_i: \text{initial size factor of service } i \\ n: \text{size of initial population} \\ N: \text{total population} \end{array} \right)$

### *b. Locality of initial user base*

Although two services start with the same size of initial population, the distribution of the population can differ. The distribution can affect evolution of social network services.

Most interesting example regarding different distribution of initial user base is the case of Facebook and MySpace. MySpace offered the service to the general public from the beginning, but

Facebook starts with college students in the dormitory. After obtaining solid user base which was locally-concentrated, Facebook started offering its service to the public. These two services evolved through different paths. Although it is not fully attributable to the different distribution of initial user base, it is worth taking the factor into consideration.

There has been studies dealing with measuring centrality of given network structure.<sup>15</sup> However, since the locality of only a certain part of given network should be measured, I would propose a new measurement as follow. (1) Compute the shortest path-length between a pair of members. (2) Sum the values of all possible pairs.

$$D_i = \sum_{j,k} p_{jk}$$

$$\left( \begin{array}{l} D_i : \text{dispersion factor of service } i \\ j, k : \text{individual } i, j \\ p_{jk} : \text{shortest path-length between } i \text{ and } j \end{array} \right)$$

Figure 8 below shows an example of the calculation.

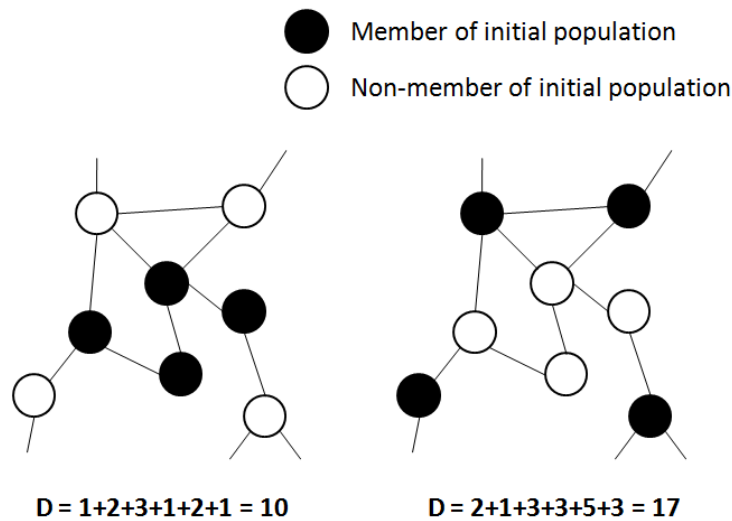


Figure 8. Example of dispersion factor calculation

<sup>15</sup> Freeman, L. C., Centrality in social networks: Conceptual clarification, Social networks, 1979

## **5. Conclusion**

Social network analysis as an academic field of study has risen since the 70s. Since the 90s, various internet-based social network services have come out to the real world. Although many studies have investigated the structural characteristics of social networks and modeled social networks, there is less literature attempting to develop a comprehensive model for describing social network services.

In this paper, social network services are economically modeled. I argue that the combination of existing research on network topology and quantified model of incentives for joining a social network service can better describe micro interactions within a social network service.

In addition to individual-level interaction modeling, initial user base of social network sites are modeled as well. Initial state of social network can be described by two factors: the size and the dispersion of initial population.

In order to solve differential equations, initial or boundary conditions are required in addition to the equation describing the system. The model of micro interactions describes dynamic part of social network services, and the model of initial user base is similar to boundary conditions. I argue that understanding these two at the same time will provide better understanding of how social network services prosper and perish.

## References

- Barabasi, & Albert, Emergence of Scaling in Random Networks, *Science* (1999)
- Boyd & Ellison, Social Network Sites: Definition, History, and Scholarship, *Journal of Computer-Mediated Communication* 13 (2008) 210-230
- Eubank et al, Modeling Disease Outbreaks in Realistic Urban Social Networks, *Nature* 429, 180 (2004)
- Freeman, Centrality in social networks: Conceptual clarification, *Social networks* (1979)
- Geroski, Models of technology diffusion, *Research Policy* 29 (2000) 603-625
- Granovetter, The Strength of Weak Ties, *The American Journal of Sociology* 78 (1973) 1360-1380
- Haythornthwaite, Social networks and Internet connectivity effects, *Information, Communication, & Society*, 8(2) (2005) 125-147
- Hoffman, Jager, & Von Eije, Social Simulation of Stock Markets: Taking It to the Next Level, *Journal of Artificial Societies and Social Simulation* vol. 10, no. 2, 7
- Li & Du, *Artificial intelligence with uncertainty*, Chapman & Hall/CRC
- Skyrms & Pemantle, A dynamic model of social network formation, *PNAS* vol. 97, no. 16, (2000)
- Wasserman & Faust, *Social Network Analysis: Methods and Applications*, New York, NY: Cambridge University Press. (1994)
- Waxman, Routing of multipoint connections. *Selected Areas in Communications. IEEE Journal on.* (1988)
- Yelle, The Learning Curve: Historical Review and Comprehensive Survey, *Decision Science*, (1979)