



# NK model-based analysis of technological trajectories: a study on the technological field of computer graphic processing systems

Ichiro Watanabe<sup>1</sup> · Soichiro Takagi<sup>2</sup>

Received: 22 April 2021 / Accepted: 4 August 2021 / Published online: 14 August 2021  
© Japan Association for Evolutionary Economics 2021

## Abstract

This paper quantitatively analyzes how the interdependence of components and the complexity of technology relates to the formation of technological trajectories. This paper uses the idea of technological trajectory and a method called a main path analysis. Technological trajectory is an idea that describes the path-dependent technological evolution process. The technological trajectories of a technological field can be represented as the main paths of patent citation networks. This paper aims to elucidate some of the determinants of the evolution of technological trajectories using main path analysis. The hypotheses are derived from a model called the NK Model. The NK model describes the respective roles of the interdependence of components and of complexity in complex adaptive systems. Using the NK model, it can be understood that technologies with an intermediate level of interdependence and technologies with an intermediate level of complexity tend to be more successful than other technologies. According to the result, the patents on the main paths of this technological field are concentrated at the intermediate level of interdependence but the patents on the main paths of this technological field are not concentrated at the intermediate level of technological complexity. Additionally, in the technological field's early stage, the interdependence values of patents that are locked-in within technological trajectories are high, whereas the same values of the later stage are low. This observation is also consistent with the idea of technological trajectories. These results suggest that the NK model is a useful tool to understand the formation of technological trajectories.

**Keywords** Technological trajectory · Technology life cycle · Patent citation network · NK model

**JEL Classification** O30 · O32 · O33 · L63

---

✉ Ichiro Watanabe  
ichrwtb@gmail.com

Extended author information available on the last page of the article

## 1 Introduction

This paper quantitatively analyzes how the interdependence of components and the complexity of technology relates to the formation of technological trajectories. Although digital technology is rapidly evolving and becoming increasingly incorporated into daily life, the understanding of the technological evolution process is still incomplete. To improve this understanding, this paper aims to contribute to the elucidation of the mechanism of the technological evolution process in a quantitative way.

This paper uses the idea of technological trajectory and a method called a main path analysis. Dosi (1982) described the technological evolution process using two notions, “technological paradigm” and “technological trajectory”. In a technological field, development paths are dependent on the technological paradigm. The technological paradigm is selected in the initial stage of the technological field’s development. Technological development must proceed according to the selected technological paradigm. The direction of technological development is limited by the technological paradigm. Dosi (1982) called these paradigm-dependent directions of development “technological trajectories”. Verspagen (2007) developed this research by introducing a quantitative method. With this method, Verspagen (2007) demonstrates that the technological trajectories of a technological field can be described as the main paths of patent citation networks. The main paths are a technological field’s major flows of knowledge that are represented as citation networks. Technological development is a selective process. While there are many possible directions of technological development, only a small fraction of these directions are realized as technological trajectories. On the other hand, the formation of the main paths is also selective. While there are many possible main paths, only a small fraction of these are realized. Therefore, technological trajectories can be represented by main paths.

In previous research, the authors examined how technological trajectories evolve within the technological field of computer graphic processing systems using main path analysis (Watanabe and Takagi 2021). However, the determinants of this evolution are still unclear. This paper aims to elucidate some of the determinants of evolution. The hypotheses are derived from a conceptual model called the NK Model. The NK model, which was originally developed in the field of evolutionary biology (Kauffman and Weinberger 1989), describes the respective roles of the interdependence of components and of complexity in complex adaptive systems. In this paper, the interdependence of components is called simply “interdependence”. The complexity of a complex adaptive system is defined by the interaction between the number of system elements and their interdependence (Fleming and Sorenson 2001; Ganco 2013, 2017). Like animals, technologies are also complex adaptive systems. They are both constituted by components that are dependent on each other. Hence, many studies have applied the NK model to technological evolution analysis. Using the NK model, the inverted *U*-shaped relationship between interdependence and the usefulness of inventors’ efforts can be understood. In the same way, using the NK model, the inverted *U*-shaped relationship between technological complexity and

the usefulness of inventors' efforts can also be understood. The inverted *U*-shaped relationship, which is presented in Fig. 1, means that an intermediate level of interdependence or complexity is optimal for the success of inventors' efforts. This relationship suggests that technologies with an intermediate level of complexity and technologies with an intermediate level of interdependence tend to be more successful than other technologies. According to this suggestion, two hypotheses can be derived: First, patents on the main paths of a technological field are concentrated at the intermediate level of interdependence. Second, patents on the main paths of a technological field are concentrated at the intermediate level of technological complexity. This paper tests these hypotheses empirically using patents within the technological field of computer graphic processing systems. First, the interdependence and the technological complexity of each patent in this technological field is calculated using the methodology of Ganco (2013). Second, this paper compares the distribution of interdependence of patents on the main paths of the technological field and the distribution of interdependence of all patents in the technological field. This paper also compares the distribution of technological complexity of patents on the main paths of the technological field and the distribution of technological complexity of all patents in the technological field. Additionally, in this paper, the change in interdependence values of patents that are locked-in within technological trajectories is analysed. This additional analysis provides deeper insight into the relationship between the idea of technological trajectory and the NK model.

The technological field of computer graphic processing systems is chosen as the target of this research. The importance of this technological field is growing together with the evolution of artificial intelligence (AI). Image recognition is an important technological field of AI. Additionally, many manufacturers use computer-aided design (CAD) software to design products. GPUs are necessary to use CAD software on PCs. The technological evolution of the technological field of computer graphic processing systems has great significance for the productivity of the manufacturing industry. Thus, examining the technological evolution of computer graphic

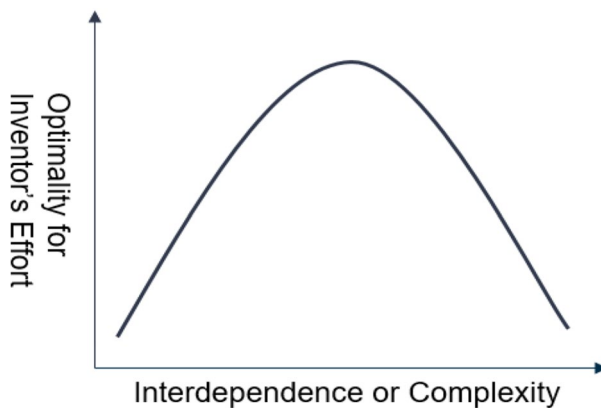


Fig. 1 Inverted *U*-shaped relationship

processing systems is particularly significant at the present time, which is why this technological field was chosen as the target of this research.

This paper consists of eight sections. After this introduction, a literature review follows. In the literature review, previous studies that are related to the idea of technological trajectories, main path analysis, and the NK model are reviewed. In Sect. 3, a brief history of the technological field of graphic processing systems is presented. In Sect. 4, hypotheses for the main analysis are presented. In Sect. 5, the data and the methodology of the main analysis are introduced. In Sect. 6, the results of the main analysis are presented. In Sect. 7, additional analysis on the change in interdependence values of patents that are locked-in is presented. Discussion about the results follows in Sect. 8.

## 2 Review of previous studies

In this section, two categories of previous studies are reviewed. First, previous studies related to the idea of technological trajectories and main path analysis are reviewed. Then, previous studies related to the NK model are reviewed.

### 2.1 Technological trajectories and main path analysis

As mentioned in Sect. 1, the idea of a technological trajectory was presented by Dosi (1982). Technological trajectories consist of cumulative and path-dependent development paths within a technological field. The idea of technological trajectories has often been used in the research field of technological evolution. Since Dosi (1982), researchers have mainly used qualitative methods to study technological evolution and technological trajectories (Possas et al. 1996; Vincenti 1994). Thus, in the research field of technological trajectories, qualitative research has been mainly cumulative. On the other hand, Verspagen (2007) advanced the research field by proposing a quantitative method to find technological trajectories within a citation network data set. Verspagen (2007) employed a method called main path analysis proposed by Hummon and Dereian (1989). Main path analysis is a method to map the major flow of knowledge within a field as citation networks. There are two steps to this method. First, every edge in the whole citation network of a field is weighed by connectivity. This weight is called traversal weight. This count represents the significance of an edge. Second, based on traversal weight, the main paths are searched by an algorithm which chains important edges in the citation network. The procedure of main path analysis will be explained in detail in Sect. 5.2. Hummon and Dereian (1989) developed this method to find the main knowledge flow within the research field of DNA studies. Verspagen (2007) argued that the method of main path analysis could be applied to find technological trajectories within a patent citation network. The main paths, which represent the major flow of knowledge within a technological field, can be considered as technological trajectories mapped as citation networks. Verspagen (2007) applied main path analysis to find technological trajectories within the technological field of fuel cells. After Verspagen (2007),

some studies employed main path analysis to find technological trajectories within the technological fields of data communication (Fontana et al. 2009) and telecommunications switching (Martinelli 2012). In addition, Barberá-Tomás et al. (2011) confirmed the validity of main path analysis as a method for studying technological evolution. In addition to these studies, Huenteler et al. (2016a) applied the main paths of patent citation networks calculated by the method of Hummon and Dereian (1989) to study the difference between technological life-cycles of solar PV and wind power. Besides, Huenteler et al. (2016b) applied the main paths of patent citation networks calculated by the method of Hummon and Dereian (1989) to study the effect of a product's design hierarchy on the evolution of technological knowledge. Main path analysis is also used to study knowledge evolution in various academic research fields. In these studies, academic paper citation network data is examined using main path analysis. For example, Yu and Sheng (2020) used main path analysis to study knowledge evolution in the research field of blockchain technology. Other studies examined the knowledge evolution of the research field of data quality, environmental innovation, IT outsourcing, text mining, data envelopment analysis, new energy vehicles, lithium iron phosphate batteries and the Internet of Things (Xiao et al. 2014; Barbieri et al. 2016; Liang et al. 2016; Jung and Lee 2020; Liu et al. 2013; Yan et al. 2018; Hung et al. 2014; Fu et al. 2019). Thus, main path analysis has been commonly used as a method to study technological and knowledge evolution. However, the method of Hummon and Dereian (1989) fails to include the edges which have large traversal counts in some cases. Liu and Lu (2012) proposed a method called key-route search to solve this problem. In this research, the authors use the key-route search method to find the main paths in the whole citation network data. After Liu and Lu (2012), additional new approaches to main path analysis have been proposed. For example, Liu and Kuan (2016) proposed a new approach to main path analysis taking into account knowledge decay.

## 2.2 NK model

The NK model, which was originally introduced by Kauffman and Weinberger (1989) in the field of evolutionary biology and discussed further by Kauffman (1993), is a conceptual model that describes the respective roles of interdependence and of complexity in complex adaptive systems. In the NK model, a complex adaptive system is described as a binary string of  $N$  components. Each digit of the string represents a component, which has two states, 0 and 1. For example, when  $N=3$ , a binary string 0 1 1 is one of eight possible combinations. The other possible combinations would be 0 0 0, 0 0 1, 0 1 0, 1 0 0, 1 0 1, 1 1 0, 1 1 1. Each combination of  $N$  components has its fitness value. Fitness means how adaptive a system or a component is to the environment. In other words, a high fitness value represents a high possibility of success of the system in an environment. The fitness value for each combination of components is calculated as the average value of the fitness values of the components. The fitness values of the components are assigned randomly. The procedure of this random assignment is determined by  $K$  values, which represent the interdependence among components of the system. For the case of  $N=3$ ,

the minimum  $K$  value is 0, which means there is no interdependence among components. The maximum  $K$  value is 2, which means each component is dependent on the two other components. For the case of  $K=0$ , a value from a uniform  $[0, 1]$  distribution is assigned to the fitness value of each binary component. In this case, the fitness value of each component does not change when the state of the other components change because the  $K$  value is 0, which means the components are not dependent on each other. Figure 2 illustrates the case of  $N=3$  and  $K=0$ . In the table of Fig. 2, all possible combinations of the binary string of 3 components are indicated in the left column. The fitness values of components 1, 2, and 3 are indicated in the middle columns as  $w_1$ ,  $w_2$ , and  $w_3$ . The fitness value for each combination of components is indicated in the right column as  $W$ . On the other hand, for the case of  $K=2$ , the fitness value assigned to each binary component depends not only on whether that component is 0 or 1 but also on whether the other two components that are interdependent with the component are 0 or 1. In this case, the fitness values of components show chaotic behaviour because a change of a component in the system (for example, a change from a component 0 to a component 1) changes the fitness of the other two components. This chaotic behaviour of the fitness values of components makes it impossible to predict the effect of a change of a component on the fitness values of the other two components. This behaviour occurs because the components are dependent on each other. This dependence is described by  $K=2$ . Kauffman (1993) describes this chaotic behaviour by assigning a fitness value to the binary components of each combination randomly from a uniform  $[0, 1]$  distribution. Figure 3 illustrates this case of  $N=3$  and  $K=2$ . In the table of Fig. 3, all possible combinations of the binary string of 3 components, the fitness values of components, and the fitness value for each combination of components are indicated in the same way as Fig. 2.

In the NK model, it is assumed that all possible combinations of  $N$  components form a fitness landscape. The fitness landscape is also an idea that was originally developed in the field of evolutionary biology by Wright (1932). Fitness landscapes are presented on the right side of Figs. 2 and 3. In a fitness landscape, agents search locally in fitness landscapes. This “search locally” means that agents change the state of one component at a time and see whether that change increases the fitness

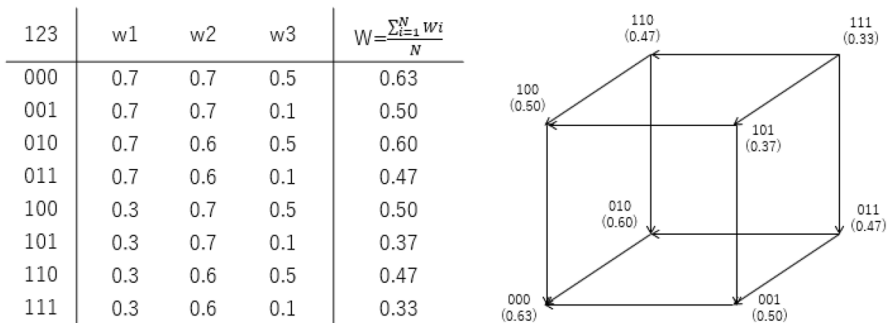
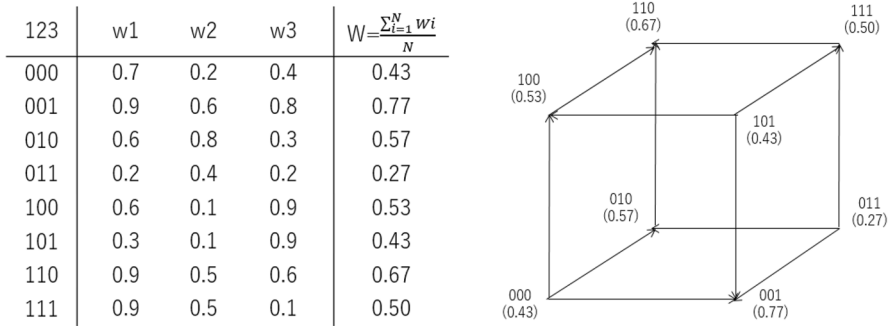


Fig. 2 The case of  $N=3, K=0$



**Fig. 3** The case of  $N=3, K=2$

value of the whole system. This local search procedure is indicated as arrows in fitness landscapes on the right side of Figs. 2 and 3. In the fitness landscape in Figs. 2 and 3, each vertex represents a combination of components. Additionally, the fitness value of each combination, which is calculated in the table, is indicated below in parentheses. The key insight from this conceptual model is that the higher the interdependence of components, the more rugged the fitness landscape. In Fig. 2, agents can reach the highest peak (the combination with the highest fitness value) in the landscape by searching locally from anywhere in the fitness landscape. On the other hand, in Fig. 3, agents cannot always reach the highest peak (the combination 001) by just searching locally. Agents sometimes become stranded at a local high peak (the combination 110) because of the ruggedness of the landscape. Thus, when  $K$  increases, it becomes difficult to find the highest peak because of the increasing ruggedness of the landscape. This means that when  $K$  increases, problem-solving becomes more difficult for agents. On the other hand, when  $K$  increases, the height of the highest peak on the landscape increases at the same time (Fleming and Sorenson 2001). This means that when  $K$  increases, the ideal fitness which agents can obtain from problem-solving also increases. According to these two effects of the change in  $K$ , it can be understood that an intermediate level of interdependence is best for the success of agents’ local search. Kauffman (1993) also mentioned a phenomenon called “complexity catastrophe”. “Complexity catastrophe” means that when the interdependence of a system is high compared to the number of components, agents tend to be stranded at a lower peak. The complexity of a complex adaptive system is represented by the ratio of  $K$  to  $N$  (Fleming and Sorenson 2001; Ganco 2013, 2017). “Complexity catastrophe” can be represented as the inverted  $U$ -shaped relationship between technological complexity value and the usefulness of agents’ efforts (Ganco 2017). In other words, “complexity catastrophe” suggests that an intermediate level of technological complexity is best for the success of agents’ local search.

The NK model has been used to understand the technological evolution process as a combinatorial search (Ethiraj and Levinthal 2004; Frenken 2000, 2006; Fleming and Sorenson 2001, 2004; Ganco 2013, 2017; Murmann and Frenken 2006; Sorenson et al. 2006; Taalbi 2017). These studies are part of a lineage of research since

Schumpeter (1934) that holds that technological novelty comes from the recombination or synthesis of existing technologies. In particular, Fleming and Sorenson (2001), Ganco (2013), and Ganco (2017) are important previous works for this paper. These papers empirically examined the validity of the NK model using patent data. Fleming and Sorenson (2001) developed a methodology to calculate the  $K$  value and technological complexity value of patents and empirically tested some hypotheses which can be derived from the NK model. The result of Fleming and Sorenson (2001) supported the hypothesis that an intermediate level of interdependence is best for the success of agents' local search. However, their result could only find partial support for the hypothesis that an intermediate level of technological complexity is best for the success of agents' local search. Ganco (2013) and Ganco (2017) developed the research of Fleming and Sorenson (2001). In the study of Fleming and Sorenson (2001), the methodology to calculate the  $K$  value and technological complexity value of patents was conducted in an inter-industry context. On the other hand, Ganco (2013) developed a methodology to calculate the  $K$  value and technological complexity value of patents in a single-industry context. Additionally, Ganco (2017) empirically tested some hypotheses which can be derived from the NK model using the calculation methodology of Ganco (2013). The result of Ganco (2017) supported the hypothesis that an intermediate level of interdependence is best for the success of agents' local search. Additionally, the result of Ganco (2017) also supported the hypothesis that an intermediate level of technological complexity is best for the success of agents' local search.

### 3 A brief history of computer graphic processing systems

In this section, the history of computer graphic processing systems is reviewed. This review is based on work by Das and Deka (2015). The technological field of computer graphic processing systems has advanced together with the evolution of the graphics processing unit (GPU). In 1999, NVIDIA introduced the term "GPU". Until this time, the term "GPU" did not exist. However, this term will be used throughout this section to ensure consistency. GPUs are designed for 3D graphics rendering calculations. The original GPU designs were based on the graphics pipeline concept. The graphics pipeline is a conceptual model that consists of several stages. Through the stages, 3D space is converted to 2D pixel space on the screen. In the early GPU hardware, only the rendering stage of the graphics pipeline was implemented. The graphics pipeline stages which are implemented in GPU hardware increased as GPU technology advanced. In 1999, the first GPUs, which implemented the whole graphics pipeline (transform, lighting, triangle setup and clipping, rendering) in their hardware were released. GeForce 256 of NVIDIA and Radeon 7500 of ATI are examples of these first true GPUs. The first graphics pipeline completely implemented in GPU hardware was called a "fixed function" pipeline because the data which was sent to the pipeline could not be modified. In 2001, NVIDIA released GeForce 3 which implemented the programmable pipeline. Using the programmable pipeline, the data can be operated while in the pipeline. The programmability of GPUs began to progress from 2001. Other examples of GPUs at this time are ATI Radeon 8500



and the Xbox of Microsoft. In 2010, NVIDIA released a GPU architecture called Fermi Architecture. This architecture was designed for general-purpose computing on graphics processing units (GPGPU), which allowed programmers to use GPU resources not only for graphics processing. Thus, the GPU hardware has advanced from a single core, fixed-function hardware pipeline implementation just for graphics to a set of programmable cores for general computing purposes.

## 4 Hypotheses

As mentioned in Sect. 2.2, two hypotheses can be derived from the NK model; (1) an intermediate level of interdependence among the components of a technology is best for the success of agents' local search, and (2) an intermediate level of technological complexity is best for the success of agents' local search. Fleming and Sorenson (2001) and Ganco (2017) used regression analysis to test these hypotheses. They checked the inverted-*U* shaped correlations between *K* value and the citation count of patents, and between technological complexity value and the citation count of patents. In these studies, the citation count of patents was used to measure the success of agents' local search. However, the citation count of patents is not the only representation of the success of agents' local search. In this research, the authors assume that the patents belonging to the main paths of a technological field represent the success of agents' local search. As mentioned in Sect. 2.1, the main paths, which can be considered as technological trajectories mapped as citation networks, represent the major flow of knowledge within a technological field. A patent belonging to the major flow of knowledge can be considered as a successful patent. Based on this assumption, two hypotheses are derived:

- (i) Patents on the main paths of a technological field are concentrated at the intermediate level of interdependence.
- (ii) Patents on the main paths of a technological field are concentrated at the intermediate level of technological complexity.

## 5 Data and methodology

In this section, the data and the methodology of the analysis are presented. The analysis aims to test the hypotheses which are presented in Sect. 4. To accomplish this goal, the following steps are taken. First, the patent citation network dataset of the technological field of computer graphic processing systems is processed using the methodology of Verspagen (2007) and Liu and Lu (2012). Second, the technological complexity and interdependence of each patent in this technological field are calculated. Third, this paper compares the distribution of interdependence of patents on the main paths of the technological field and the distribution of interdependence of all patents in the technological field. This paper also compares the distribution of

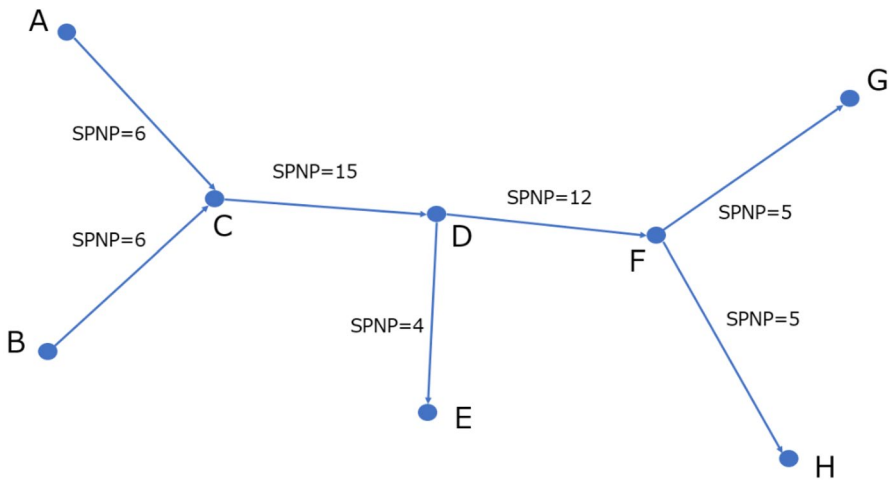
technological complexity of patents on the main paths of the technological field and the distribution of technological complexity of all patents in the technological field.

## 5.1 Patent data

In this section, the data and the methodology which are used for the analysis are introduced. The US Patent Office database is used to obtain the entire patent citation network data in the technological field of computer graphic processing systems. This field is defined by the technological classes of US Patent Classification (USPC) under Class 345/501. There are eight subclasses (345/502, 345/503, 345/504, 345/505, 345/506, 345/519, 345/520, 345/522) under this class. Class 345/501 is for the technological field of the “Computer graphic processing system”. According to the class definition, patents of “subject matter comprising apparatus or a method for processing or manipulating data for presentation by a computer prior to use with or in a specific display system” (USPC class numbers and titles, Class 345/501) are classified under Class 345/501. There are many more patents that are essential for the evolution of computer graphic processing systems. For example, patents which are classified as Class 382 are about image analysis, which is a subject that is strongly related to computer graphic processing systems. However, in this research, the patents which are not included in the classes under Class 345/501 are not examined to keep the data manageable. In addition, the citations that are examined in this research are citations within the classes under Class 345/501. The US Patent Office online database called PatentsView covers the patents which are published since 1975. In this study, patents from 1975 to 2015 were obtained from PatentsView. The history of computer graphic processing systems started in the 1970s, so the scope of the dataset is adequate for this research. The number of patents that are collected from PatentsView is 4032. After collecting patents, a citation network data set of the technological field of computer graphic processing systems is created. Python and its network analysis package NetworkX are used to create this citation network data set. In the citation network data set, every node represents a patent, and every directed edge represents a citation. The citation network data set in this research contains 4032 nodes and 13,147 edges. Every edge is directed to a citing patent from a cited patent according to the flow of knowledge. For example, in Fig. 4, an edge is directed to node C from node A. This edge represents a relationship in which patent C cites patent A. Patent citation networks are always directed acyclic graphs (DAGs) because no patents cite patents that are newer than them. In addition, some patents are never cited but cite others. Such patents become sink nodes in the network. Sink nodes are called “endpoints” in this research. At the same time, some patents are cited but cite nothing in the citation network. Such patents become source nodes of the network. Source nodes are called “startpoints” in this research.

## 5.2 Main path analysis

In the main path analysis method, every edge in the citation network data is first weighted according to its position in the network. The weight of edges is called



**Fig. 4** Calculating SPNP values

the “traversal count”. The search path count method is used to weigh every edge in an acyclic network. The term “search path” means a route that connects a pair of nodes in the network. Every search path is a sequence of directed edges. For example, in Fig. 4, a search path A-C-D-F connects the node A to the node F. There are two search path count methods, search path link count (SPLC) and search path node pair (SPNP). Search path link count (SPLC) is a method proposed by Hummon and Dereian (1989). In this method, every edge is weighed by counting how often the edge lies on all possible search paths. Hummon and Dereian (1989) imply that the SPLC method contains search paths whose origins are intermediate nodes or search paths whose destinations are intermediate nodes. However, the method can also be considered to contain only search paths whose origins are startpoints and whose destinations are also endpoints. In this research, the SPLC method is considered to contain only search paths whose origins are startpoints and destinations are also endpoints. Hummon and Dereian (1989) also proposed another method to weigh edges. This method is called the search path node pair (SPNP). The edge D-F connects four nodes (A, B, C, D) to its destination, for example, the node G. At the same time, the edge D-F connects three nodes (F, G, H) to its origin, for example, the node A. The SPNP value of the edge D-F is calculated by multiplying these numbers. Thus, the SPNP value of the edge D-F is  $3 \times 4 = 12$ . This number represents how many pairs of nodes the edge D-F connects. Both SPLC and SPNP weigh nodes which are more responsible for connecting other nodes. As Verspagen (2007) and Fontana et al. (2009) mentioned, the result is not very different between these two methods. In this research, the SPNP method is used following Verspagen (2007) and Fontana et al. (2009). After finishing weighing edges, the following algorithm proposed by Hummon and Dereian (1989) is adopted to define main paths within a network. The algorithm below is created in reference to Verspagen (2007).

- (i) For each startpoint in the network, pick the outward edge(s) that has(ve) the highest SPNP value among all the edges going out from the startpoint. Each startpoint is assigned a value equal to the value of the edge(s) with the highest SPNP value.
- (ii) Select the startpoint(s) with the highest assigned value. This is the startpoint(s) of the main paths.
- (iii) Take the target(s) (citing patent) of the edge(s) identified in the previous step.
- (iv) From the target(s) identified in the previous step, pick (again) the outward edge that has the maximum SPNP value among all outward edges from this node and add this edge to the main paths. If some edges have the same maximum SPNP value, add all these edges to the main paths. If (all) these edge(s) point to an endpoint of the network, exit the algorithm, otherwise go back to Step (iii) and continue.

The main paths which can be found by this algorithm are called the local main paths. This algorithm suffers from the limitation that sometimes it does not include the edges which have large traversal counts. To avoid this problem, Liu and Lu (2012) invented a new method called key-route search. Key-route search guarantees that the local key-route main paths, which are calculated by the method, contain edges with the highest traversal counts. According to Liu and Lu (2012), the key-route search procedure is as follows. The algorithm below is created in reference to Liu and Lu (2012).

- (i) Select the key-route, which is the links that have the highest traversal count.
- (ii) Search forward from the end node of the key-route until a sink is hit.
- (iii) Search backward from the start node of the key-route until a source is hit.

“Search forward” is the same as steps (iii) and (iv) of Hummon and Dereian (1989)’s method presented previously. “Search backward” represents searching the roots of the edges using steps (iii) and (iv) of Hummon and Dereian (1989). The local key-route main paths are calculated by this procedure. The local key-route main paths within the network of Fig. 4 are presented by thick lines in Fig. 5. The authors use the key-route search method to find the main paths in the whole citation network data. A network analysis software called Pajek is used to conduct the main path analysis. The procedure for analyzing main paths in Pajek is explained in detail in de Nooy et al. (2018).

In this research, the main path analysis was repeated for a sequence of periods to obtain a series of “snapshots” of the main paths of each year from 1985 to 2015. In total, 31 “snapshots” of main paths are obtained. All patents which belong to these 31 “snapshots” are grouped as patents on the main paths.

### 5.3 Calculation of the interdependence and complexity of each patent

In this paper, the  $K$  values, which represent the interdependence among components, and the technological complexity values of patents are calculated using the

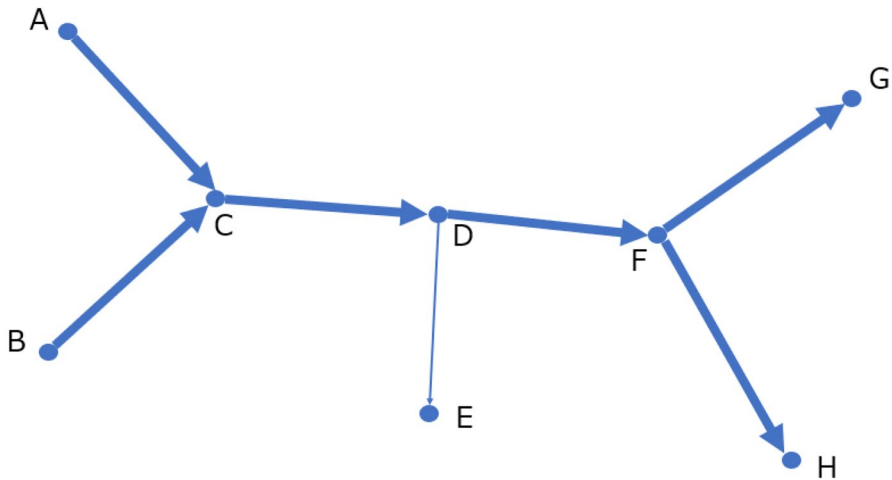


Fig. 5 The local key-route main paths of the network presented in Fig. 4

methodology of Ganco (2013). As referenced in Sect. 2.2, Fleming and Sorenson (2001) originally developed the methodology to calculate the  $K$  values and the technological complexity values of patents. This methodology was conducted in an inter-industry context. Later Ganco (2013) developed a methodology to calculate the  $K$  values and the technological complexity values of patents within a single-industry. Since the target technological field of this research is single, the methodology of Ganco (2013) is suitable for this research. In the methodology of Ganco (2013), the  $K$  values of patents are calculated in several steps. First, the  $K_i$  values of the components are calculated. The patent subclasses to which a patent belongs are considered as the knowledge components of which the patent consists. The  $K_i$  value of the components (patent subclasses) of patent  $l$  is described as:

$$K_i = \sum_{j \in L_i} \frac{\text{count of patents in subclasses } i \text{ and } j}{\text{count of patents in subclass } i}$$

where  $j$  belongs to all patent subclasses except  $i$ . Subclass  $i$  and  $j$  are any patent subclasses to which at least one patent that is within the target data of this research belongs. Second, the  $K$  values of each patent are calculated. The  $K_l$  value of patent  $l$  is described as:

$$K_l = \frac{1}{\text{count of subclasses of patent } l} \sum_{i \in l} K_i$$

Ganco described the logic behind this methodology as follows:

The key idea behind the measure is that when two underlying functions (represented by patent subclasses) are coupled, components belonging to these classes are more likely to occur in a single invention. If the functions  $A$  and  $B$  are highly coupled, if component  $a$  is classified in patent subclass  $A$ ,  $a \in A$ , and if component  $b$  is in subclass  $B$ ,  $b \in B$ , then one is more likely to see subclasses  $a$  and  $b$  in a single

invention. In other words, high interdependence between  $A$  and  $B$  implies that whenever an inventor solves a problem related to one of these functions, she/he needs to redesign or include the coupled function as well, and the components optimizing these functions are likely to be observed together in a patent. Similarly, if the patent improves the architecture of multiple functions, all components that correspond to these functions are likely to be coupled to the architecture. On the other hand, if  $A$  and  $B$  are independent with respect to each other,  $A$  is likely to be combined with other subclasses without  $B$  being present. (Ganco 2013: 676).

$N_l$ , which is the number of components that constitute patent  $l$ , is the number of patent subclasses to which patent  $l$  belongs. The technological complexity value of patent  $l$  is described as the ratio of  $K_l$  to  $N_l$ ,  $\frac{K_l}{N_l}$  (Fleming and Sorenson 2001; Ganco 2013, 2017).

## 6 Result

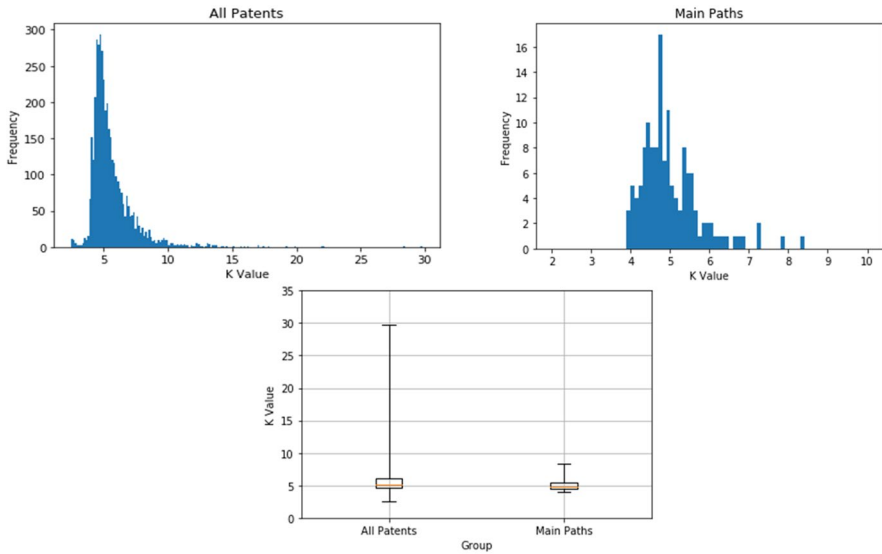
In this section, the distribution of  $K$  values and technological complexity values of patents on the main paths of the technological field of computer graphic processing systems, and the distribution of  $K$  values and technological complexity values of all patents in this technological field are compared.  $K$  values and technological complexity values of patents are calculated using the methodology presented in Sect. 5.3.

The basic statistics of  $K$  values and technological complexity values of each group are presented in Table 1. From Table 1, it can be said that the  $K$  values of patents on the main paths are concentrated in a narrower range than the  $K$  values of all patents in this technological field. On the other hand, the difference between the range of technological complexity values of patents on the main paths, and the range of the technological complexity values of all patents that in this technological field is not as large as in the case of the  $K$  values.

The histogram on the top-left side of Fig. 6 presents the distribution of  $K$  values of all patents in the technological field. The histogram on the top-right side of Fig. 6 presents the distribution of  $K$  values of patents on the main paths. The boxplot in Fig. 6 presents the comparison of distributions of  $K$  values of each group. This boxplot shows that the patents on the main paths of a technological field are

**Table 1** Basic statistics of  $K$  values and technological complexity values

	$K$ , All	$K$ , Main Paths	Complexity, All	Complexity, Main Paths
Size	4032	139	4032	139
Min	2.4842	3.9685	0.2978	0.4804
Max	29.8060	8.3868	5.8104	4.3932
Median	5.1036	4.8207	1.2676	1.2707
Mean	5.5727	5.0169	1.5120	1.4638
Variance	2.9860	0.5827	0.7189	0.5593

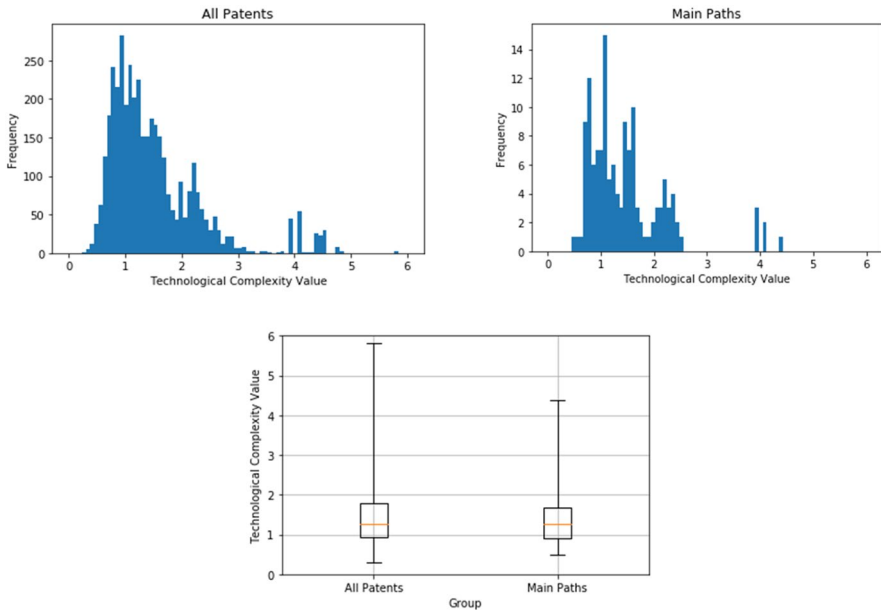


**Fig. 6** Distributions and Boxplot of  $K$  values

concentrated at the intermediate level of interdependence. The variance of the  $K$  values of all patents in the technological field is 2.9860. On the other hand, the variance of the  $K$  values of all patents on the main paths is 0.5827. To test the difference in the variances between the two groups, the Levene test was used. The p-value of the test was less than 0.001. According to Fig. 6 and the result of the Levene test, it can be said that the patents on the main paths of this technological field are concentrated at the intermediate level of interdependence.

The histogram on the top-left side of Fig. 7 presents the distribution of technological complexity values of all patents in the technological field. The histogram on the top-right side of Fig. 7 presents the distribution of technological complexity values of patents on the main paths. The boxplot in Fig. 7 presents the comparison of distributions of the technological complexity values of each group. This boxplot shows that the patents on the main paths of a technological field are slightly concentrated at the intermediate level of technological complexity. The variance of the technological complexity values of all patents in the technological field is 0.7189. On the other hand, the variance of the technological complexity values of all patents on the main paths is 0.5593. To test the difference of the variances between the two groups, the Levene test was used. The p-value of the test was 0.2968. According to Fig. 7 and the result of the Levene test, it cannot be said that the patents on the main paths of this technological field are concentrated at the intermediate level of technological complexity.

According to the result presented, Hypothesis (i) is supported and Hypothesis (ii) is not supported. In other words, the patents on the main paths of this technological field are concentrated at the intermediate level of interdependence but the patents on the main paths of this technological field are not concentrated at the intermediate



**Fig. 7** Distributions and Boxplot of technological complexity values

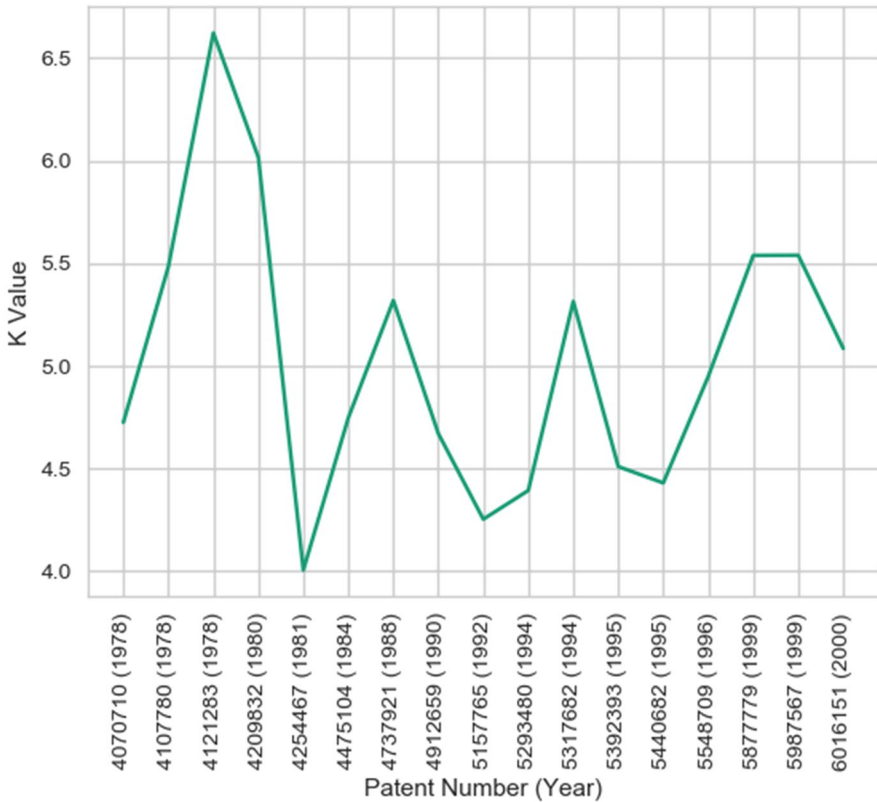
level of technological complexity. Ganco (2017) empirically revealed that there is an inverted *U*-shaped relationship between technological complexity and the usefulness of inventors' efforts in a single-industry context. The result of this paper is inconsistent with the result of Ganco (2017).

## 7 Change in the *K* values of locked-in patents over time

In this section, the change in *K* values, which represent the interdependence of the components, of patents that are locked-in will be examined to obtain additional implications. Watanabe and Takagi (2021) mentioned that all patents observed on the main paths three times consecutively at 5-year intervals did not drop out from the main paths in the long term. They also mentioned that this observation is consistent with the technological lock-in process. According to them, all patents observed on the main paths 11 times consecutively at 1-year intervals are considered to be patents that are locked-in in the technological field of computer graphic processing systems. In this section, the change in the *K* values of these locked-in patents is analysed. These locked-in patents are extracted from the “snapshots” of main paths, which are calculated in Sect. 5.2.

Figure 8 presents the change in the *K* values of patents that are locked-in within this technological field and Table 2 presents the details of the locked-in patents. The highest *K* values are observed at the initial stage of this technological field (1978–1980, Patent No. 4121283 and 4209832) and the *K* values became stable within the lower range after the initial stage. This observation is consistent





**Fig. 8** Change in the  $K$  values of locked-in patents over time

with the idea of technological trajectory. Dosi (1982) mentioned that there are two kinds of technological progress: extraordinary breakthrough innovations and normal technological progress. Breakthrough innovations create a new technological field and set the technological paradigm of the technological field. Subsequent normal technological progress is accumulated along the technological trajectory of the technological field. Normal technological progress tries to extend the possibility of breakthrough innovations as much as possible and when further extension becomes impossible, innovators search for a new paradigm (Arthur, 2010; Dosi, 1982). The two patents in Fig. 8 with the highest  $K$  values can be considered as the breakthrough innovations in this technological field. There are two reasons for this. First, as mentioned before, these patents are observed at the technological field's initial stage. Second, as explained in Sect. 2.2, high  $K$  values mean that the difficulty of problem-solving is high. At the same time, also as explained in Sect. 2.2, the ideal fitness values that innovators can obtain from solving these problems are higher than those from solving problems with lower  $K$  values. Breakthrough innovations are relatively rare, which means such types of innovations' problems are difficult to solve. At the same time, breakthrough

**Table 2** Patents presented in Fig. 8

Patent No.	Year	Assignee	Title
4070710	1978	NuGraphics, Inc	Raster scan display apparatus for dynamically viewing image elements stored in a random access memory array
4107780	1978	National research development corporation	Display apparatus
4121283	1978	Cromemco Inc	Interface device for encoding a digital image for a CRT display
4209832	1980	Chrysler Corporation	Computer-generated display for a fire control combat simulator
4254467	1981	Xerox Corporation	Vector to raster processor
4475104	1984	Lexidata Corporation	Three-dimensional display system
4737921	1988	Dynamic Digital Displays, Inc	Three dimensional medical image display system
4912659	1990	International Business Machines Corporation	Parallel surface processing system for graphics display
5157765	1992	International Business Machines Corporation	Method and apparatus for pipelined parallel rasterization
5293480	1994	AT&T Bell Laboratories	High resolution graphics system architecture
5317682	1994	International Business Machines Corporation	Parametric curve evaluation method and apparatus for a computer graphics display system
5392393	1995	Sun Microsystems, Inc	Architecture for a high performance three-dimensional graphics accelerator
5440682	1995	Sun Microsystems, Inc	Draw processor for a high performance three-dimensional graphic accelerator
5548709	1996	Silicon Graphics, Inc	Apparatus and method for integrating texture memory and interpolation logic in a computer system
5877779	1999	Sun Microsystems, Inc	Method and apparatus for efficient rendering of three-dimensional scenes
5987567	1999	Apple Computer, Inc	System and method for caching texture map information
6016151	2000	NeoMagic Corp	3D triangle rendering by texture hardware and color software using simultaneous triangle-walking and interpolation for parallel operation

innovations must have some universality to maintain the paradigm for a long time. This universality can be interpreted as high fitness. The two patents in Fig. 8 with the highest  $K$  values can be considered to have high problem-solving difficulty. However, they have belonged to the technological trajectory of this technological field for a long time. This observation means that the innovators of these patents solved difficult problems and realised high fitness. In Sect. 6, many patents with high  $K$  values which do not belong to technological trajectories are observed. It can be said that the innovators of these patents were not able to realise high fitness by difficult problem-solving. Thus, these features of the two patents with high  $K$  values (Patent No. 4121283 and 4209832) are consistent with the features of breakthrough innovations. On the other hand, the subsequent normal technological progress occurs frequently, which means such types of innovations' problems are not as difficult to solve as those of breakthrough innovations. Additionally, normal technological progress does not need to have as high universality as breakthrough innovations. These features are consistent with the features of innovations with low  $K$  values. Thus subsequent patents in Fig. 8 with low  $K$  values (Patent No. 4254467 ~) can be considered to be patents of normal technological progress. Thus, the observation presented in this section is consistent with the idea of technological trajectories. In future research, the authors will conduct qualitative research to verify whether or not Patent 4121283 and Patent 4209832 are breakthrough innovations of this technological field. Additionally, the generality of the observation will be tested in future research.

## 8 Discussion

According to the result presented in Sect. 6, Hypothesis (i) is supported and Hypothesis (ii) is not supported. In other words, the patents on the main paths of this technological field are concentrated at the intermediate level of interdependence but the patents on the main paths of this technological field are not concentrated at the intermediate level of technological complexity. Ganco (2017) empirically revealed that, in a single-industry context, there is an inverted  $U$ -shaped relationship between technological complexity and the usefulness of inventors' efforts. The result of this paper is inconsistent with the result of Ganco (2017). The authors will examine the causes of this inconsistency in future research. Additionally, according to the result presented in Sect. 7, in the technological field's early stage, the  $K$  values of patents that are locked-in within technological trajectories are high, whereas the same values in the later stage are low. This observation is also consistent with the idea of technological trajectories. The result which is presented in Sect. 6 and Sect. 7 suggests that interdependence of components affects the evolution of technological trajectories and the NK model is a useful tool to understand the formation of technological trajectories. However, these results are not enough to say that interdependence is one of the determinants of the evolution of technological trajectories. The authors will do further research to verify whether or not interdependence is one of the determinants of the evolution of technological trajectories. Additionally, this paper has the limitation of having only examined patents in the technological field of computer

graphic processing systems. The generality of the result should be tested in future research.

A theoretical implication is derived from this research: patents on main paths tend to have an intermediate level of interdependence. This paper revealed this previously unknown feature of patents on main paths in the case of patents in the technological field of computer graphic processing systems. The authors believe that this will be confirmed as a proposition applicable to technologies in general as research results in other fields become available.

An implication for science and technology policymaking is also derived from this research: investments in technologies that do not have an intermediate level of interdependence should be avoided because their probability of being a part of a technological trajectory seems to be relatively low. This implication is derived from the analysis of patents in the technological field of computer graphic processing systems. It is shown that patents without intermediate interdependence do not belong to the main path in this technological field. The generality of this conclusion should be tested in future research.

**Acknowledgements** The authors benefited from comments by and discussion with Professor Osamu Sakura of the Interfaculty Initiative in Information Studies at the University of Tokyo and Professor Makoto Nirei of the Faculty of Economics at the University of Tokyo. The authors also benefited from discussion at the 19th International Conference of the Japan Economic Policy Association.

#### Declarations

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest. This article does not contain any studies involving human participants performed by any of the authors.

## References

- Arthur WB (2010) *The nature of technology: what it is and how it evolves*. The Free Press (Simon & Schuster), New York
- Barberá-Tomás D, Jiménez-Sáez F, Castelló-Molina I (2011) Mapping the importance of the real world: the validity of connectivity analysis of patent citations networks. *Res Policy* 40(3):473–486
- Barbieri N, Ghisetti C, Gilli M, Marin G, Nicolli F (2016) A survey of the literature on environmental innovation based on main path analysis: a survey of the literature on environmental innovation. *J Econ Surv* 30(3):596–623
- Das PK, Deka GC (2015) History and Evolution of GPU Architecture. In: Deka GC, Siddesh GM, Srinivasa KG (eds) *Emerging Research Surrounding Power Consumption and Performance Issues in Utility Computing*. IGI Global, Pennsylvania, pp 109–135
- de Nooy W, Mrvar A, Batagelj V (2018) *Exploratory Social Network Analysis with Pajek*. Rev. and expanded, 3rd edn. Cambridge University Press, Cambridge
- Dosi G (1982) Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. *Res Policy* 11(3):147–162
- Ethiraj SK, Levinthal D (2004) Modularity and innovation in complex systems. *Manage Sci* 50(2):159–173
- Fleming L, Sorenson O (2001) Technology as a complex adaptive system: evidence from patent data. *Res Policy* 30(7):1019–1039

- Fleming L, Sorenson O (2004) Science as a map in technological search. *Strateg Manag J* 25(89):909–928
- Fontana R, Nuvolari A, Verspagen B (2009) Mapping technological trajectories as patent citation networks. an application to data communication standards. *Econ Innov New Technol* 18(4):311–336
- Frenken K (2000) A complexity approach to innovation networks. the case of the aircraft industry (1909–1997). *Res Policy* 29(2):257–272
- Frenken K (2006) Innovation, evolution and complexity theory. E. Elgar, Cheltenham
- Fu H, Wang M, Li P, Jiang S, Wei Hu, Guo X, Cao M (2019) Tracing knowledge development trajectories of the internet of things domain: a main path analysis. *IEEE Trans Industr Inf* 15(12):6531–6540
- Ganco M (2013) Cutting the Gordian knot: the effect of knowledge complexity on employee mobility and entrepreneurship. *Strateg Manag J* 34(6):666–686
- Ganco M (2017) NK model as a representation of innovative search. *Res Policy* 46(10):1783–1800
- Huenteler J, Schmidt TS, Ossenbrink J, Hoffmann VH (2016a) Technology life-cycles in the energy sector — technological characteristics and the role of deployment for innovation. *Technol Forecast Soc Chang* 104:102–121
- Huenteler J, Ossenbrink J, Schmidt TS, Hoffmann VH (2016b) How a product's design hierarchy shapes the evolution of technological knowledge—evidence from patent-citation networks in wind power. *Res Policy* 45(6):1195–1217
- Hummon NP, Dereian P (1989) Connectivity in a citation network: the development of DNA theory. *Social Networks* 11(1):39–63
- Hung S-C, Liu JS, Lu LYY, Tseng Y-C (2014) Technological change in lithium iron phosphate battery: the key-route main path analysis. *Scientometrics* 100(1):97–120
- Jung H, Lee BG (2020) "Research Trends in Text Mining: Semantic Network and Main Path Analysis of Selected Journals. *Expert Syst Appl* 162:113851
- Kauffman SA (1993) The origins of order: self organization and selection in evolution. Oxford University Press, New York
- Kauffman SA, Weinberger ED (1989) The NK model of rugged fitness landscapes and its application to maturation of the immune response. *J Theor Biol* 141(2):211–245
- Liang H, Wang J-J, Xue Y, Cui X (2016) IT Outsourcing research from 1992 to 2013: a literature review based on main path analysis. *Inform Manag* 53(2):227–251
- Liu JS, Kuan C-H (2016) A new approach for main path analysis: decay in knowledge diffusion. *J Am Soc Inf Sci* 67(2):465–476
- Liu JS, Lu LYY (2012) An integrated approach for main path analysis: development of the hirsch index as an example. *J Am Soc Inform Sci Technol* 63(3):528–542
- Liu JS, Lu LYY, Wen-Min Lu, Lin BGY (2013) Data envelopment analysis 1978–2010: a citation-based literature survey. *Omega (oxford)* 41(1):3–15
- Martinelli A (2012) An emerging paradigm or just another trajectory? understanding the nature of technological changes using engineering heuristics in the telecommunications switching industry. *Res Policy* 41(2):414–429
- Murmann JP, Frenken K (2006) Toward a systematic framework for research on dominant designs, technological innovations, and industrial change. *Res Policy* 35(7):925–952
- Possas ML, Salles-Filho S, da Silveira JoséMaria (1996) An evolutionary approach to technological innovation in agriculture: some preliminary remarks. *Res Policy* 25(6):933–945
- Schumpeter JA (1934) *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard University Press, Cambridge
- Sorenson O, Rivkin JW, Fleming L (2006) Complexity, networks and knowledge flow. *Res Policy* 35(7):994–1017
- Taalbi J (2017) What drives innovation? evidence from economic history. *Res Policy* 46(8):1437–1453
- Verspagen B (2007) Mapping technological trajectories as patent citation networks: a study on the history of fuel cell research. *Adv Complex Syst* 10(1):93–115
- Vincenti WG (1994) The retractable airplane landing gear and the northrop "Anomaly": variation-selection and the shaping of technology. *Technol Cult* 35(1):1–33
- Watanabe I, Takagi S (2021) Technological trajectory analysis of patent citation networks: examining the technological evolution of computer graphic processing systems. *Rev Socionetw Strateg* 15(1):1–25
- Wright S (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: *Proceedings of the sixth international congress on genetics*, vol. 1, pp 356–366

- Xiao Yu, Lu LYY, Liu JS, Zhou Z (2014) Knowledge diffusion path analysis of data quality literature: a main path analysis. *J Informet* 8(3):594–605
- Yan J, Tseng F-M, Lu LYY (2018) Developmental trajectories of new energy vehicle research in economic management: main path analysis. *Technol Forecast Soc Chang* 137:168–181
- Yu D, Sheng L (2020) Knowledge diffusion paths of blockchain domain: the main path analysis. *Scientometrics* 125(1):471–497

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Ichiro Watanabe<sup>1</sup>  · Soichiro Takagi<sup>2</sup>

Soichiro Takagi  
takagi.soichiro@mail.u-tokyo.ac.jp

- <sup>1</sup> Graduate School of Interdisciplinary Information Studies, The University of Tokyo, Tokyo, Japan
- <sup>2</sup> Interfaculty Initiative in Information Studies, The University of Tokyo, Tokyo, Japan