

**Toward valuable prediction of ERP diffusion in North American automotive industry: a simulation based approach**

Adnène Hajji, Robert Pellerin, Ali Gharbi, Pierre Majorique Léger, Gilbert Babin

**Abstract**

Recent researches suggest that the decision to adopt an ERP system has implications outside of firm boundaries, and is likely to impact the adoption decision from stakeholders in the same industrial network. Being able to consider such phenomena, to reproduce the whole dynamic of the diffusion process and to validate empirical evidences can be determinant in the success of a new product introduction. This paper is intended to propose and validate a propagation model for the diffusion of enterprise resource planning (ERP) systems among the North American automotive business network. Based on empirical evidence from the adoption of ERP by top 50 firms in the automotive sector from 1994 to 2005, a micro quantitative diffusion model is proposed. A simulation-based approach is employed to estimate the parameters of the proposed model to represent faithfully the observed diffusion phenomenon. The results show a robust diffusion model with a mean error less than 4% and up to 92% of similarity between the real ERP adopters data gathered through publicly available secondary databases and those obtained via simulations. Finally, the conducted validation experimentations show that when we progress in time the cumulative historical data minimizes the effort of estimation and future adopters' prediction is more effective. Discussion about the usefulness of the proposed model for ERP systems editors to test different strategies for market penetration is also conducted.

**Keywords** ERP, diffusion model, network, secondary data, simulation.

## 1- Introduction

Enterprise Resource Planning (ERP) systems are commercial software packages that enable the integration of their complete range of business processes and functions across an organisation ([Davenport 1998](#); [Markus et al. 2000](#)). The vast majority of large organizations have adopted ERP systems, and, only a few large software manufacturers dominate the enterprise systems market ([Hestermann et al. 2012](#)).

ERP adoption decision and acceptance has widely been investigated through the lenses of individual level factors that influence the usage of the ERP system ([Hwang 2005](#); [Kwak et al. 2012](#)). However, recent researches suggest that the decision to adopt an ERP system has implications outside of firm boundaries, and is likely to impact the adoption decision from stakeholders in the same industrial network. Building on institutional theory ([Dimaggio and Powell, 1983](#)), [Aubert et al. \(2012\)](#) find that stakeholders with strong ties are linked to have an influence on organizations making an ERP adoption decision. [Pellerin et al. \(2007\)](#) find homophily in ERP adoption decision between Canadian firms with board interlocks. [Millaire et al. \(2009\)](#) report that this network, influencing the diffusion of ERP, is particularly significant in the early stages of adoption. Taken together, these results are in line with the call to consider the decision for organization to adopt an ERP from a diffusion of innovation perspective ([Light et al. 2004](#)).

In this paper, we build upon previous researches to propose a simulation-based tool to validate the extension of the network-based diffusion model of ERP systems, developed upon Valente model ([Valente 1995](#)), and proposed in [Millaire et al. \(2009\)](#). The proposed model describe the dynamic diffusion of an organizational-based innovation in an industrial network and allow to test various network intervention penetration strategies ([Hajji et al. 2012](#); [Choi et al. 2012](#) ; [Li et al. 2013](#)). We argue that taking into account an agent level decision involved in a given network is necessary to model the network diffusion process. This necessity has been the center of attention in several research studies, such as ([Chatterjee and Eliashberg 1990](#); [Park and Hahn 2000](#); [Van Eck et al. 2011](#)).

The automotive sector was chosen, as a first candidate, to validate our approach due to the large amount of data available relevant to this industrial network and the strategic nature of this industry. Real network and ERP adoption data were collected through various publicly secondary data sources. A simulation-based approach, combined with a genetic algorithm, is employed to validate and refine the proposed model, therefore allowing an accurate representation of the diffusion phenomenon. It is shown that such proposed diffusion model represents faithfully the reality under a mean error of diffusion lesser than 4%, and a good percentage of similarity between the real ERP adopters and those found via simulations. Moreover, the conducted validation experimentations show that, when we progress in time, the cumulative historical data minimizes the effort of estimation, and future adopters' predictions are more effective.

The remainder of the paper is organized as follows: Section 2 presents a review of network based diffusion of industrial technology. The proposed diffusion model is presented in Section 3. Section 4 than present our data collection methodology followed the presentation of the simulation-based optimization approach used to validate and

experiment the proposed model. A discussion of the experimental results obtained is reported in Section 6. The paper concludes in Section 7 by underlining our contribution and future research activities in this area.

## **2- Network based diffusion of industrial technology: Literature review**

Diffusion of innovation in networks and social contagion has become an important topic of research in organizational studies and the interest for external pressure and network-based influence on organizational decision with regards to industrial technologies goes beyond the ERP literature. Network-based models of diffusion of innovations have been widely studied in many related organizational studies (Strang and Soul 1998; [Valente 1995](#)). Researches in organization sciences provide evidence that organizations tend to imitate each other by adopting new technologies or practices to gain legitimacy, due to network externalities or simply by managerial fashion (Abrahamson 1996; [Davis 1991](#); [Deephouse 1996](#)). For example, studies have investigated how opinion leadership and social contagions affect the adoption of new products ([Iyengar et al., 2011](#)). Abrahamson and Rosenkoft (1997) argue that social network structure plays a crucial role in explaining the bandwagon effect which underlies the diffusion of innovation with strong network externalities.

Early research on network diffusion of innovations has identified social networks as an important factor to consider ([Coleman et al. 1957](#)). The use of sociometry (DeNooy et al. 2005) was shown to be useful in understanding how an innovation flows within the community and how certain network structural variables influence the diffusion process (Trunda and Westarp 1998). A network is defined as the pattern of relationships, ideas, decisions or support that exists among members of a social system (Trunda and Westarp 1998). Two approaches exist to study how relationships influence the diffusion of innovations through the members of a given network, namely, the relational and the structural approach. The relational approach considers the focal actor's adoption in light of the behaviour of those to which it is directly connected. Thus, for a given actor, direct contact with an influential leader might greatly influence its choice of system (Trunda and Westarp 1998; [Shaw-Ching et al. 2005](#)). Structural network models acknowledge that the overall structure of the network, as well as a given actor's position in it, influence that actor's behaviour and subsequent performance. For example, a given actor may adopt an innovation because of its prior adoption by a highly central and visible actor, even though the two actors may have no direct contact with each other ([Valente 1995](#)).

It has been previously argued that ERP are software technologies susceptible to network externalities, meaning that the perception of utility of this system is tied to the adoption network ([Aubert et al. 2012](#)). Not all software products generate the same externalities as they depend on the level of economies of scale in production and distribution; learning effect; knowledge sharing; and technological compatibility ([Fichman 2004](#)). In the case of an ERP system, its network externalities come from various elements. First of all, these business software applications have been developed, over time, to account for business processes across a large number of industries. The more installation of a specific system, the richer the software becomes and can be more easily implemented in various companies. Moreover, early adopters can also benefit from these additional functionalities when the system is upgraded. Also, software, with a larger install-based,

may benefit from a larger number of consultants available in order to deploy the software, leading to broader expertise on the software at a lower cost.

From a theoretical standpoint, previous studies in network diffusion of ERP have built upon institutional theory ([DiMaggio and Powell 1983](#)). This theory asserts that institutions are defined by the socially shared behaviours. Individuals are pressured to reproduce institutional patterns which reinforce the legitimacy of the behaviours and leads to isomorphism. Previous research shows that this isomorphism extends to interorganizational behaviors, specifically in the existence of network externalities.

[Aubert et al \(2012\)](#) provide evidence that neighbouring companies linked with strong ties can have an influence on organisations making ERP adoption decisions. Past decisions made by major trading partners in the chemicals and allied products sectors have a significant influence on the decision to adopt an ERP system for a given organisation. [Millaire et al \(2009\)](#) also shows that two types of network exposure appear to have an influence on the diffusion of ERP in industrial networks: a) the network exposure which takes place from the direct relationship ties of an actor in the network and b) the industry exposure which represents the indirect pressure from the community. Direct exposure is defined as the number of firms in a business relationship with a particular company that has already adopted an enterprise system. As for indirect exposure to the network, the definition is simply the exposure of a company to all the enterprises in its industry sector that have already adopted a system.

The following sections will present the proposed model, and a simulation-based approach in order to optimize some parameters governing the model, and to reproduce more faithfully the considered diffusion processes.

### **3- The proposed diffusion model**

The network structure and its heterogeneity are key elements that must be considered to model faithfully the reality of an industrial network-based diffusion process. In their work largely cited in the scientific literature, [Chatterjee and Eliashberg \(1990\)](#) stated that aggregate diffusion models where no deliberate decision is involved limit their consideration of the determinants of adoption at the individual level. Thus, to address the unrealistic assumption of population homogeneity by aggregate (macro-level) diffusion models, micro-level diffusion models have been developed ([Wedel and Kamakura 2000](#)). Micro-model diffusion models often use classical utility and attitude models and attempt to model effects arising from innovation uncertainty or differences in individual perception structure ([Bohlmann et al. 2010](#)). [Chatterjee and Eliashberg \(1990\)](#) developed an analytical model of the innovation diffusion process for a high involvement service. This model explicitly considers the determinants of adoption at the individual level and allows for heterogeneity with respect to these determinants across the population. They adopted a micro-modeling approach that postulates individual adoption times as an explicit function of the characteristics of potential adopters. [Bohlmann et al. \(2010\)](#) have also mentioned the advantage of the micro-modeling approach when considering the market network heterogeneity.

The proposed model is an extension of Valente's model (Shaw-Ching et al. 2005) and previous results from (Coleman et al. 1957; Pellerin et al. 2007). This model takes into account both direct and indirect network externalities of enterprise system adoption and is developed based on the premise that firms are part of a complex web of interconnected business ties that directly or indirectly connect the firm to their partners and peers.

$$\log\left(\frac{\Pr(\lambda_t = 1)}{1 - \Pr(\lambda_t = 1)}\right) = \alpha + \beta_1 Sales_t + \beta_2 IE_{it} + \beta_3 NE_{it} \quad (1)$$

$$NE_i = \frac{\sum_{j \in v_i} (w_{ij} \lambda_j)}{\sum_{j \in v_i} w_{ij}} \quad (2)$$

$$IE_i = \frac{\sum_j (n_{ij} \lambda_j)}{\sum_j n_{ij}} \quad (3)$$

where  $\beta_k$  are parameter estimates for vectors of the associate effect,  $Sales_t$  is vector of sales,  $\omega$  is the social network matrix,  $v_{ij}$  is the set of companies linked to company  $i$ ,  $\lambda$  is vector of adoptions and  $\eta_{ij}$  indicates whether companies  $i$  and  $j$  are in the same industry sector.  $NE_i$  and  $IE_i$  are respectively the network and industry exposure of company  $i$ .  $Sales_t$  is the vector of companies' sales at time  $t$ . This component allows linking the adoption decision to the sales performance of a given company over time.

Recall that the model and the propositions made in (Coleman et al. 1957) were tested using the social network analysis approach, as several authors have applied in the past (Shaw-Ching et al. 2005). DeNooy et al. (2005) define "social network analysis" as a method to detect and interpret the patterns in a social network. It studies a network of entities, paired based on the varied nature and the dynamic effect they have on each other.

As detailed in Section 2, although the macro model has been very useful to differentiate and validate the importance of the aforementioned effects, they present limitations when considerations at the individual level should be taken.

In what follows, based on the aggregate model (1) – (3) and using a micro-modelling approach, we propose a micro model (4) – (6) which allows individual-level variables affecting adoption timing of a product in a given network. The proposed model considers explicitly the adoption decision based on an individual level evaluation of the innovation and the previously detailed network effects.

Potential adopters are assumed to evaluate the ERP system along two attributes, network effect and perceived benefits. Based on these attributes a utility function, based on the price of the ERP and its perceived value on the market, will allow an enterprise belonging to the network to make decision over time to adopt or not the ERP system.

The proposed model states the probability of adoption as follows:

$$P(\lambda_{it} = 1) = P(\lambda_{it-1} = 1) + P(V_{it} > C_t | \lambda_{it-1} = 0) \quad (4)$$

such as

$$\sum_{\forall t} \lambda_{it} \leq 1 \quad (5)$$

$$V_{it} = a + \beta_1 Sales_i + \beta_2 IE_{it} + \beta_3 NE_{it} \quad (6)$$

where:

$t$  : period

$i$  : enterprise number

$\lambda_{it}$  : node status at time  $t$  (0 = non adoption state, 1 adoption state)

$V_{it}$  : ERP systems' value for the enterprise  $i$  at time  $t$

$a$  : base value of an ERP system for all the enterprises (the perceived benefits)

$C_t$  : adoption cost of the system

Consequently, the node status is a function of the system's value given by the enterprise with respect to the system's cost. The system's value varies with respect to the enterprise's environment as it is a function of the exposition of the system in the network. The more an enterprise is surrounded by adopters, the more its value increases.

The ERP systems' value are governed by three unknown parameters  $\beta_k, k=1-3$  which should be estimated to reproduce a given diffusion process. Once estimated the challenge is to test whether the model can predict the future diffusion process with a good level of confidence. This issue is addressed in the next sections.

#### **4-Sociometric data collection and construction of industry network**

Social influence can be tested using the network analysis method only if the period covered by data collection is long enough to see the effect of the network on the adoption (Valente 1995). Consequently, we opted for a longitudinal data collection; we collected data on nodes (*who?*), ties (*how?*), technology type (*what?*) and time of adoption (*when?*).

Figure 1 illustrates the data collection process and social network development. To collect the necessary data, four sources of data were mainly used, for the accuracy of their data or for the volume of potential evidence: Compustat, PRNewswire, SEC-EDGAR and Harte-Hanks.

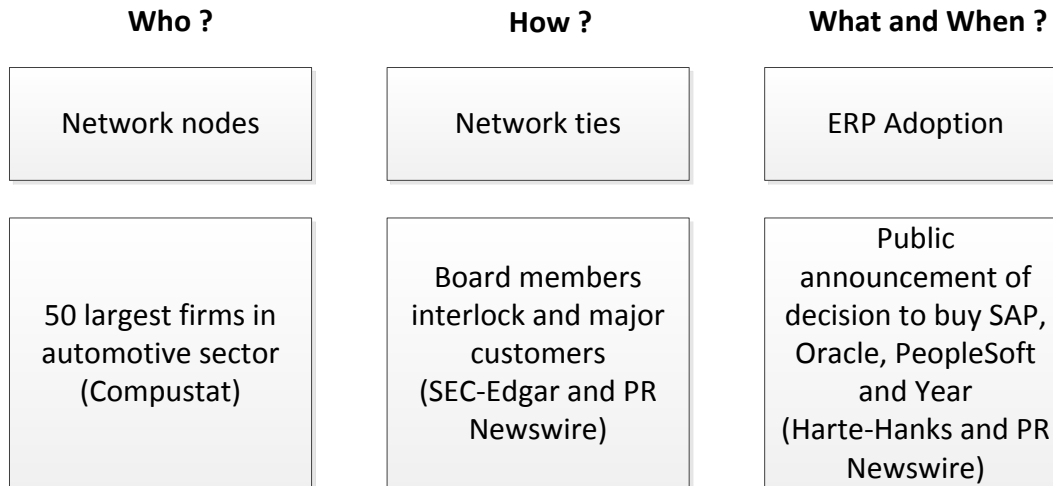


Figure 1 : Sociometric Data Collection

#### 4.1 Network nodes (who ?)

To collect the nodes, we identified the top 50 largest public firms in the North-American automotive industry for each year in the period from 1995 to 2004. The choice of this time interval was motivated by the fact that the decision to adopt an enterprise system can often take many years ([Wybo et al. 2005](#)). Compustat was used to extract this list of public company from the following automotive related SIC codes : 3011, 3537, 3711, 3713, 3714, 3715, 3716, 3751 and 5013. Due to attrition, a total of 54 unique companies are part of the total sample. These firms account for 99.46 % of the revenues in the 9 above SIC codes.

#### 4.2 Network ties

To identify the network ties between the 54 identified companies, two data sources were used. First, we extracted from the annual reports (10-K) and from the Definitive Proxy Statement (DEF 14A) of each US company from the Securities and Exchange Commission (SEC) electronic data repository (EDGAR). For Canadian company, this information was extracted from SEDAR. The first name, last name, age and tenure of each directors on every of the 12 years was collected from these reports. The year of birth was used to distinguish homonyms. Data was triangulated with the Directory of Directors from the Financial Post. Missing value from EDGAR and SEDAR were obtained by contacting directly with the department of shareholder relations of the respective companies.

The other source of ties used in this research are the supply chain partners. Customers representing more than 10% of a firm revenues, which must be mandatorily disclosed to the SEC. Using EDGAR and SEDAR, we extracted this information when available.

To enrich the list of customer, we used PR Newswire, available through Thomson-Dialog. This source was mainly used to find evidence of business relationship between the firms selected. To do this, we developed a Java based software to connect via to PR Newswire API to automate the data extraction process. Every firm was queried against



the 53 others for the period from 1995 to 2004. If the two keywords were found in the same press release, it was automatically saved in a text file. 40,585 press releases were extracted.

#### 4.3 ERP adoption (what and when ?)

Finally, ERP adoption was triangulated from two different sources. First, technological profile of the 54 firms was acquired from the CI Technology database provided by Harte-Hanks. This database contains technological profile of more than 750,000 firms in 25 countries. It contains information about systems, servers, network infrastructure and computer equipment for each company. Data is acquired and updated by Harte-Harke via an annual survey. 32 out of 50 firms were identified as having an ERP system. In other words, this database mostly contained which ERP was adopted, but not when it was adopted.

To enrich the ERP adoption data and to identify when each firm adopted their ERP, we again turned to PR Newswire. It is industry practice that the software vendor create a press release when a firm decides to purchase an ERP system. We used the same Java based software to connect to PRNewswire to automate the extraction of press releases. Press released were manually inspected to find evidence of adoption.

For companies for which no adoption data was found, we manually queried Google, Google News Archives and ABI/Informs (Proquest) with the name of the company and each of the major ERP provider. The website of ERP providers was also searched to complete the data collection.

Data was finally validated by an industry expert to confirm the exactness of our data. 32 firms out of 54 were found to have adopted an ERP systems between 1995 and 2004.

#### 4.4 Construction of industry network

A java based tool was developed to construct industry network in a format amenable to analysis, using the data previously obtained.

We refer the reader to Millaire et al. (2009), Hajji et al. (2012) and Aubert et al. (2012) for detailed discussions of the obtained results. Mainly, these results suggest a mix of influences from both exposures (direct and indirect). Consistent with Valente's demonstration of threshold and critical mass, enterprise system adoption is very well rooted in the network diffusion of innovation. The contribution of network exposure and industry exposure differs based on the time in the product life cycle and the category of firms looking to acquire a system.

### **5- Simulation based approach and implementation**

In order to bring an approach which could be easily applied to represent and predict the diffusion phenomena in a given network, the descriptive capacities of discrete/continuous event simulation models are combined with analytical models and genetic algorithms. A block diagram of the resulting control approach is depicted in Figure 2.



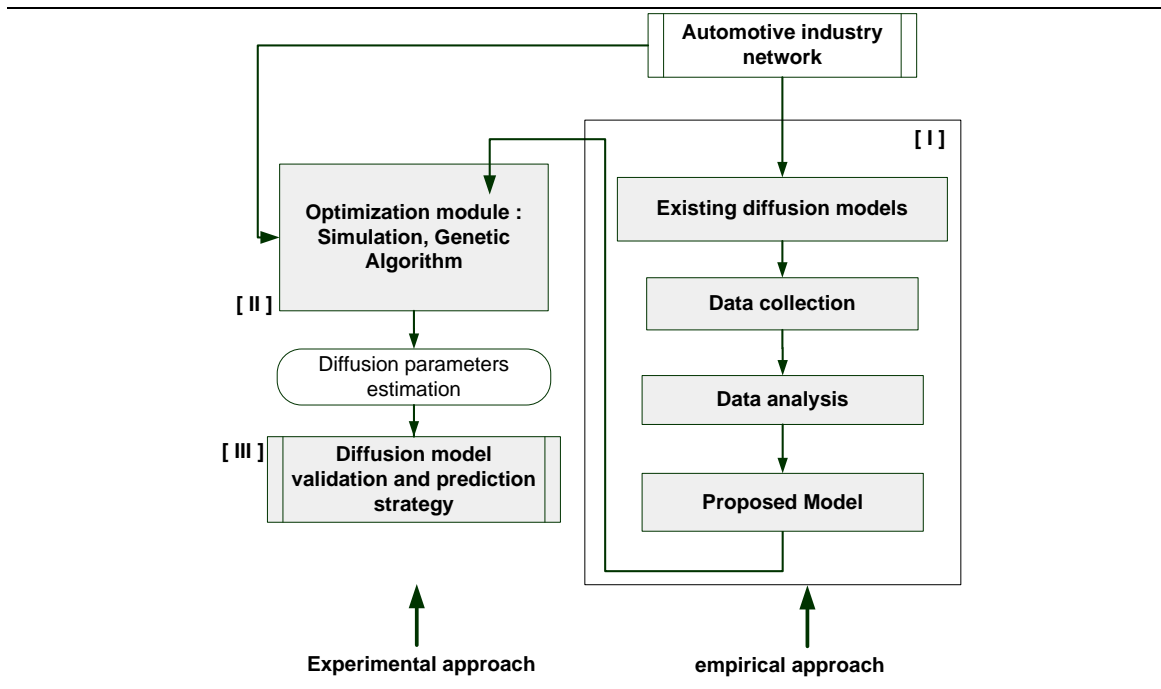


Figure 2: implementation approach

- I. The first part of the approach consists in developing the quantitative model to express the ERP diffusion phenomena. This model was presented in Section 3 and will govern the following steps.
- II. The second part consists on building an optimization module supporting the quantitative parameters governing the diffusion model. This module links a parameterized simulation model with a genetic algorithm making it possible to run a genetic algorithm search process for the best solution (minimize the gap with real diffusion). The search process is detailed in Figure 3. It consists on running the genetic algorithm with respect to its stopping rule and evaluating each desired configuration through the simulation model.

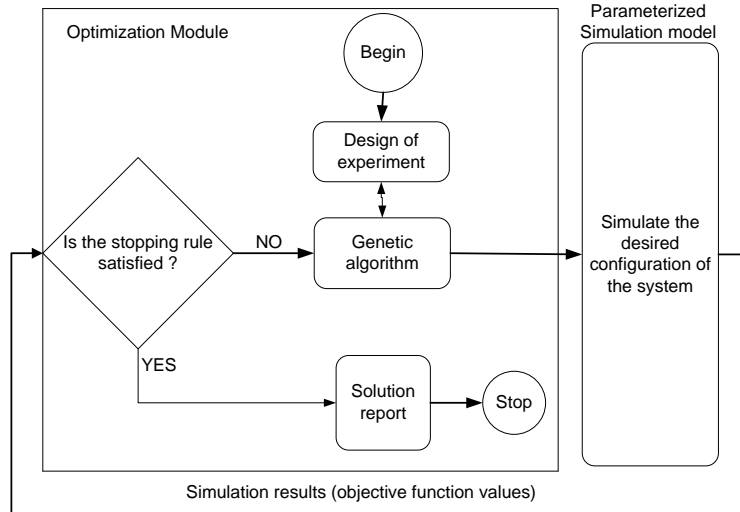


Figure 3: optimization module

III. The third part is the resulting validated diffusion model allowing offline experimentations to predict future adopters and to test different strategies for market penetration for example.

### 5.1 Simulation model

The simulation model is built to describe the dynamics of the diffusion process governed by the model defined previously and parameterized by the aforementioned parameters  $\beta_k$  (Section 3). These factors are considered as input of such a model and the related incurred diffusion curve and adopters are defined as its output. The combined discrete/continuous parameterized simulation model is developed using the Visual SLAM language (Pritsker et O'reilly 1999).

The Visual SLAM portion is composed of various networks describing specific tasks (data exchange with Genetic algorithm, random generation of adopters, status update and time advance, etc...). The simulation ends when current simulation period reaches the defined simulation horizon. Figure 4 shows a bloc diagram representation of the simulation model.

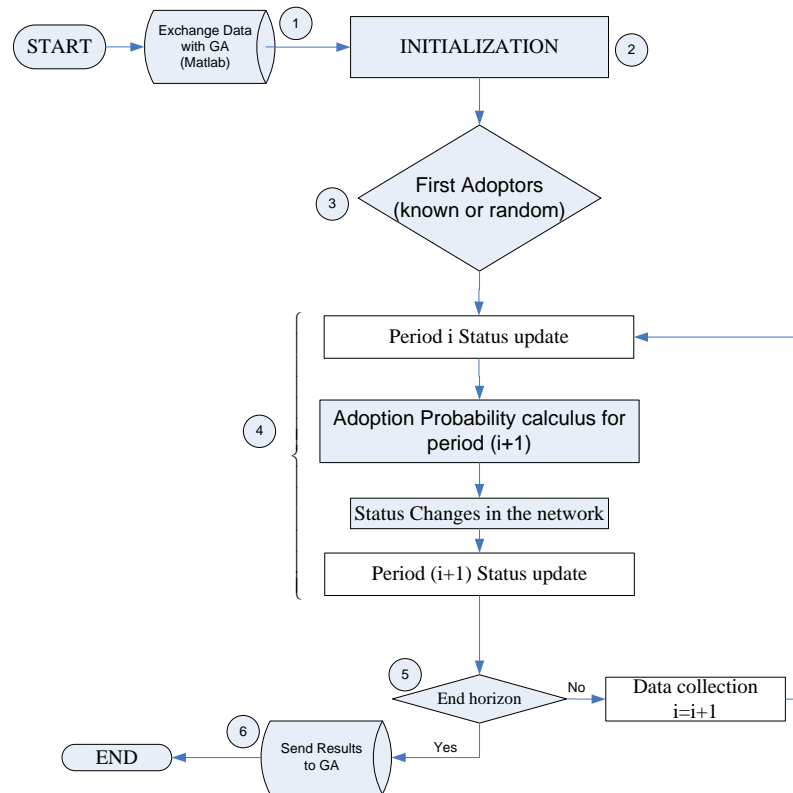


Figure 4: simulation model

The simulation model, presented in Figure 4, allows reproducing, with a series of discrete events, the continuous evolution of an enterprise status with respect to the adoption of an ERP system for a fixed number of periods. The status change of an enterprise (i.e., adoption vs. non-adoption) is governed by the diffusion model (4)-(6).

1-2) The Exchange data block read the parameters of each individual of the population set by the Genetic Algorithm. To run the model the INITIALIZATION block sets these values and other parameters defining the system (e.g., the initial network structure, simulation horizon ...).

3) The third block sets the first adopters in the network. If known, a simple update of their status is conducted. Otherwise, a random set of adopters depending on a fixed penetration rate is generated by the simulation engine.

4) The four subsequent blocks represent the core of the simulation model. They are governed by the diffusion model (4)-(6). Based on the adopters and the network structure of a given period, a prediction of the next adopters and an update of the whole diffusion process are conducted.

5) The end horizon block is used to end the simulation process or to go on with the subsequent iteration.

6) The Send Results block writes the diffusion curve (adopters in each period of the horizon) of each individual of the genetic algorithm population in an external file. This file being available to the genetic algorithm program reiterates the optimization process.

## 5.2 Genetic algorithm

In terms of an optimization problem, the genetic algorithm approach is summarized as follows. At any given point in time, the genetic algorithm generates a population of possible candidate solutions. Initially, the population size is chosen at random. However this choice typically depends on the characteristics of the problem. Each population component is a string entity of chromosome (e.g., (0,1) bit string) which represents a possible solution to the problem. The population components are evaluated based on a given fitness function. Highly fit population components are given the chance to reproduce through a crossover process with other highly fit population elements by exchanging pieces of their genetic information. This process produces “offspring” or new solutions to the optimization problem based upon the high-performance characteristics of the parents. Premature loss of important information by randomly altering bits within a chromosome is prevented by a mutation process. This procedure continues until a satisfactory solution is achieved.

## 5.3 Optimization module implementation

When genetic algorithms are employed to carry out an optimization process, each point in the solution space is characterized by a chromosome. Each position in the chromosome characterizes a decision alternative. The optimization process starts with a random sample of the solution space. Each point of this sample is sent to the simulation model to evaluate his objective function (i.e., diffusion curve in terms of adopters and similarity with real diffusion of earlier periods). After that, based on the objective function values of each point, a *selection* rule is employed to generate a new population. Chromosomes of this population are then *crossover* following fixed rules to construct a new generation. Another decision to be taken consists on a possible introduction of mutation in the process. This is done by random sampling with a given probability, commonly chosen as small value (less than 5 %).

Regarding the search process stopping rule, many of them are used in the literature. A common rule consists on stopping the algorithm when the generated populations become equivalent or when we find that there is no significant improvement. Another rule consists on running the simulation for a fixed number of runs.

To set the optimization module, 3 steps are required. Building a parameterized simulation model; building a genetic algorithm program and connecting these two sub-units to allow the communication between them.

## 5.4 Genetic algorithm: MATLAB Toolbox

The main data structures in the GA toolbox ([Chipperfield et al. 1994](#)) are chromosomes, phenotypes, objective function values and fitness values. The chromosome structure stores an entire population in a single matrix of size  $N_{ind} \times L_{ind}$ , where,  $N_{ind}$  is the number of individuals and  $L_{ind}$  is the length of the chromosome structure. Phenotypes are stored in

a matrix of dimension  $N_{ind} \times N_{var}$  where,  $N_{var}$  is the number of decision variables. A  $N_{ind} \times N_{obj}$  matrix stores the objective function values, where  $N_{obj}$  is the number of objectives. Finally, the fitness values are stored in a vector of length  $N_{ind}$ . In all of these data structures, each row corresponds to a particular individual.

The GA toolbox uses MATLAB matrix functions to build a set of versatile routines for implementing a wide range of genetic algorithm methods. In this section we outline the major procedures of the GA Toolbox and especially those used in our program.

1. Population representation and initialisation: the GA Toolbox supports binary, integer and floating-point chromosome representations.
2. Fitness assignment: the fitness function transforms the raw objective function values into non-negative figures of merit for each individual. The Toolbox supports the offsetting and scaling method of (Goldberg 1989) and the linear-ranking algorithm of (Baker 1985).
3. Selection functions: available routines include roulette wheel selection (Goldberg 1989) and stochastic universal sampling (Baker 1987).
4. Crossover operators: the crossover routines recombine pairs of individuals with given probability to produce offspring. Single-point, double-point (Baker 1987) and shuffle crossover (Caruana et al. 1989) as well as a general multi-point (Syswerda 1989) crossover routine are provided.
5. Mutation operators: Binary and integer mutation are performed by the routine `mut`. Real-values mutation is available using the breeder GA mutation function, `mutbga`.

The following steps summarize the employed Genetic Algorithm:

1. Population representation and initialisation: binary representation with «  $N_{ind}$  » the number of individuals and « `Preci` » the precision of the binary representation.
2. Fitness: the linear-ranking method of (Baker 1985).
3. Selection: stochastic universal sampling of (Baker 1987). The technique needs to fix a ratio « `GGAP` » of the best elements to keep.
4. Crossover: Single-point (Baker 1987) with crossover probability « `Pc` ».
5. Mutation: binary mutation with probability  $P_m = 1/L_{ind}$ ,  $L_{ind}$  is the length of the chromosome structure equal  $L_{ind} = \text{Preci} \times N_{var}$ .

Let « `MaxGen` » be the maximum number of generation if the stopping algorithm rule is fixed following this criteria.

## 6. Parameters estimation: results and discussion

The network population is composed of the top 54 firms in the North American auto sector between 1995 and 2004. The following figure shows the cumulative adoption over the period from 1994 to 2005 (Léger et al. 2008). The same data were used in this study with regards to number of adopters and the year of adoption.

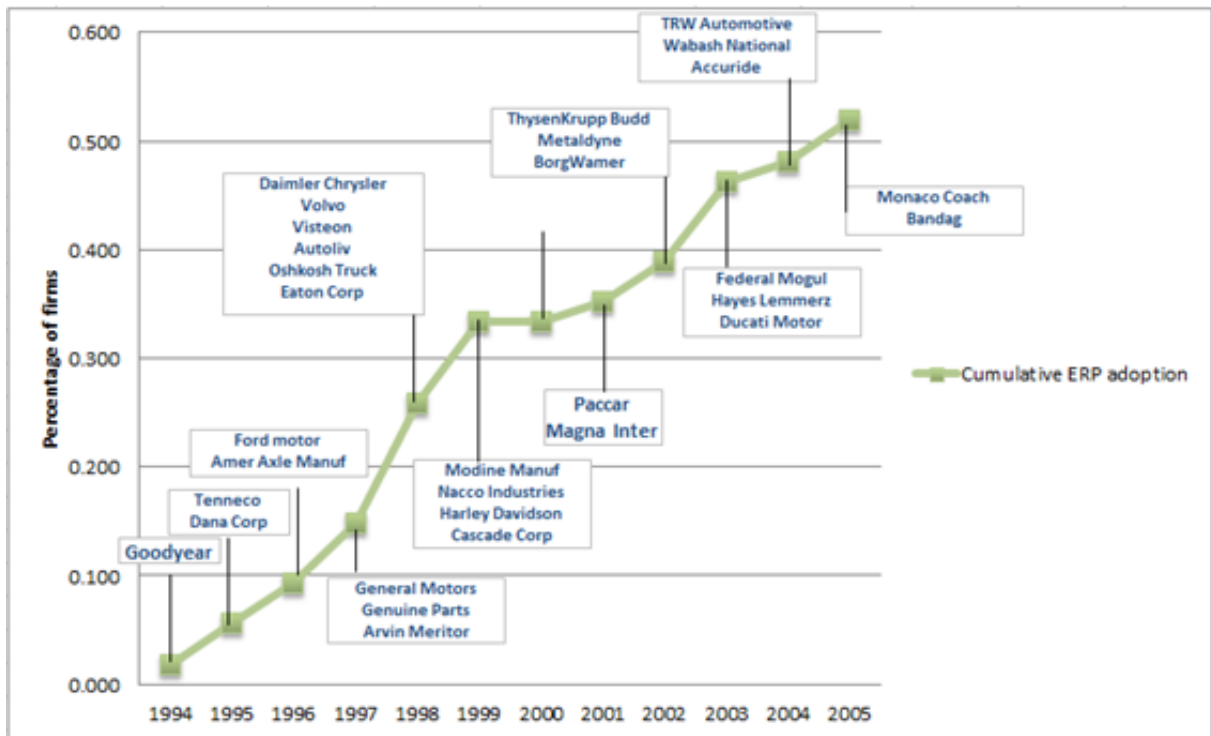


Figure 5: ERP adoption in the North American auto industry

Basically two goals are achieved. The first one consists in running the simulation and optimization engine to check the realism of the micro model and the simulation tool in representing the reality of ERP diffusion. Two measures are analysed in this part: the similarity between real versus experimental adopters and the relative error between real and experimental diffusion curves. The second goal consists in examining the benefit to be drawn from a limited diffusion history in order to make a projection for future years. In this part, the simulation and optimization engine is run only for a limited number of years and the resulting diffusion model is kept for the rest of the horizon. These two goals are detailed in the following sections.

### 6.1 Micro model validation

Starting from 1996 with the real adopters, the simulation and optimization tool is launched to find the best parameters of the proposed model (4)-(6) with the objective to minimize the gap between real and experimental diffusion curves. The optimal parameters are given in table 1. Note that from 1997 to 2005 the diffusion process is exclusively supported by the simulation model. Thus, no adjustment, to update the real adopters from year to year, have been conducted.

Table 1: Optimal diffusion model parameters

$a$	$\beta_1$	$\beta_2$	$\beta_3$
0,086005	0,016422	0,32568	0,2037

Table 2 summarizes the obtained results showing a mean relative error equal to 3,8% and a percentage of similarity equal to 67%. The real and experimental diffusion curves are illustrated in Figure 6. It is interesting to note that apart the gap observed in 1997 which can be explained by the aleativity in diffusion process at early stages, it is the observed error at 2000 which retain the attention. Such a gap is explained by Y2K bug where no new adoption has been noted (see table 2, real adopters). Although the proposed model shows a very promising description of the diffusion process, we claim that random unpredictable market events should be taken into consideration when generalizing the findings.

Table 2: Experimental Vs. Real diffusion

	Real adopters	Experimental adopters	Relative error	similarity	% similarity
1996	5	5	0,00%	5	100%
1997	8	9	12,50%	6	75%
1998	14	14	0,00%	9	64%
1999	18	17	5,56%	10	56%
2000	18	20	11,11%	11	61%
2001	19	20	5,26%	12	63%
2002	21	21	0,00%	13	62%
2003	25	25	0,00%	16	64%
2004	26	26	0,00%	17	65%
2005	28	27	3,57%	17	61%
Mean			3,8 %		67 %

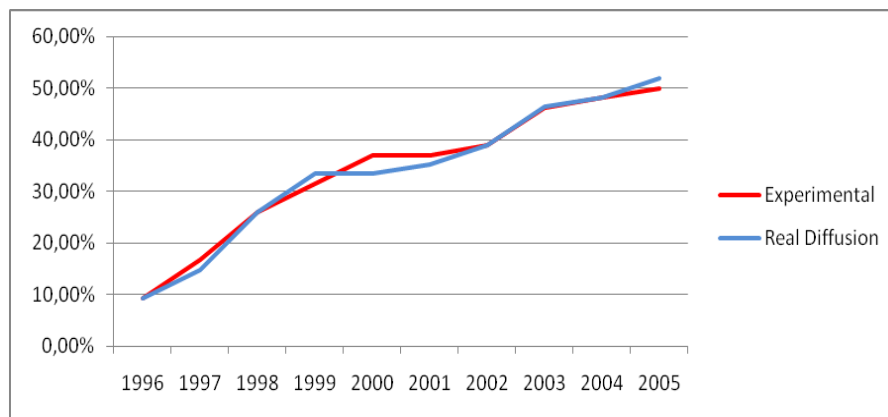


Figure 6: diffusion curve



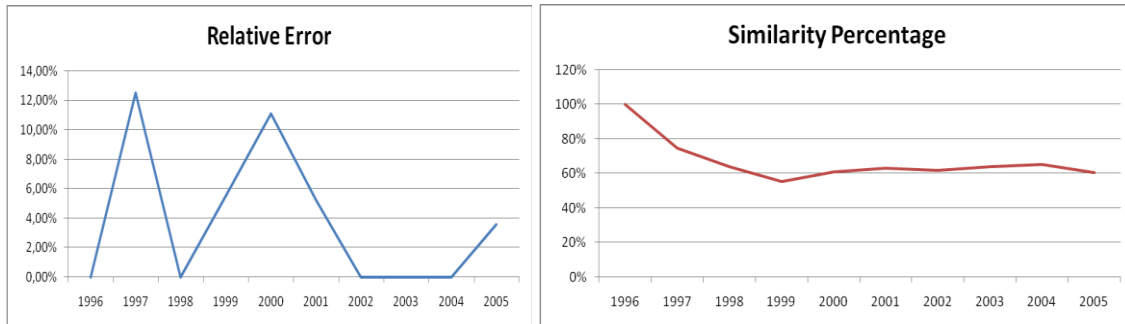


Figure 7: deviation and similarity measures

Regarding the similarity between the real adopters and those found via experimentations between 1997 and 2005, the obtained results show a mean percentage of similarity equal to 67 % (see Table 2 and Figure 7). It is interesting to note that considering only the real adopters at 1996 we were able to reach 67 % of real companies who have adopted the ERP system. This issue will be detailed in the next section to show that with a larger historical data regarding the real adopters we are able to improve our prediction for the future diffusion process with a better percentage of similarity.

## 6.2 Historical-based diffusion prediction

As discussed earlier, the main objective of our study is to provide a robust model and a powerful simulation based tool allowing an effective projection of the diffusion process and the potential adopters. The validation conducted has shown good results and fitting measures. To go forward, the same approach is used to make a projection of the future from the known available date. In other expression, we propose to proceed following four steps:

- 1- Consider the available data of a selected number of years between 1996 and 2005,
- 2- Run the optimization process to find the best parameters of the diffusion model in the considered period,
- 3- Keep the same model and launch the simulation for the reminder of the horizon.
- 4- Evaluate the relative error and the percentage of similarity between the real and experimental diffusion process.

Considering the adopters at 1996 and following the aforementioned steps has led to the diffusion curve given by Figure 6. The percentage of similarity between the real and experimental adopters are equal to 63 %, the mean relative error is equal to 12,42%. The obtained results seem to be very promising if we note that it is a projection of 9 years in the future based on only one year of real adopters historical data. This optimism is supported by the experimentations conducted with respectively 2, 3, 4 and 5 years of historical data. The obtained results are given by Figures 8 to 10 and show a very close prediction with very good percentages of similarity (see Table 3 and Figure 11). The summary table 3 show that considering 5 years of historical data, we are able to predict the next 5 years with a mean relative error of about 2% and a percentage of similarity that

exceeds 90%. These findings are very promising and confirm the robustness of the proposed model and the simulation based tool to represent faithfully the diffusion process and to help expecting the potential adopters.

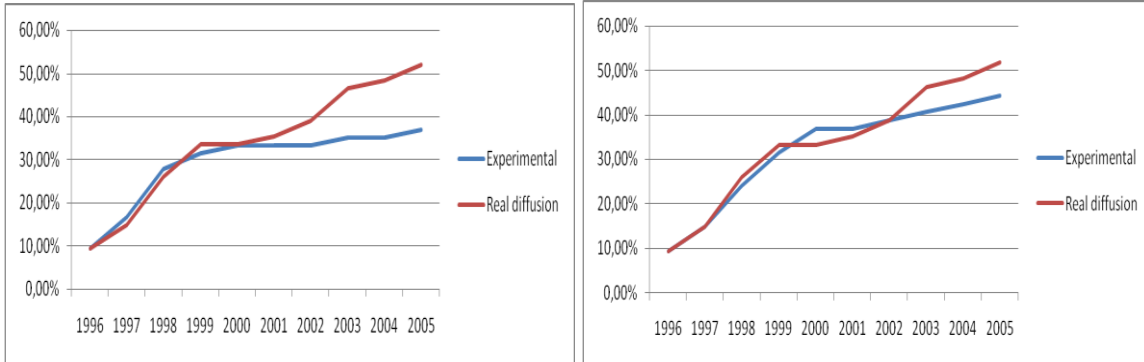


Figure 8: diffusion curves considering 1996 history (left) and 1996-97 history (right)

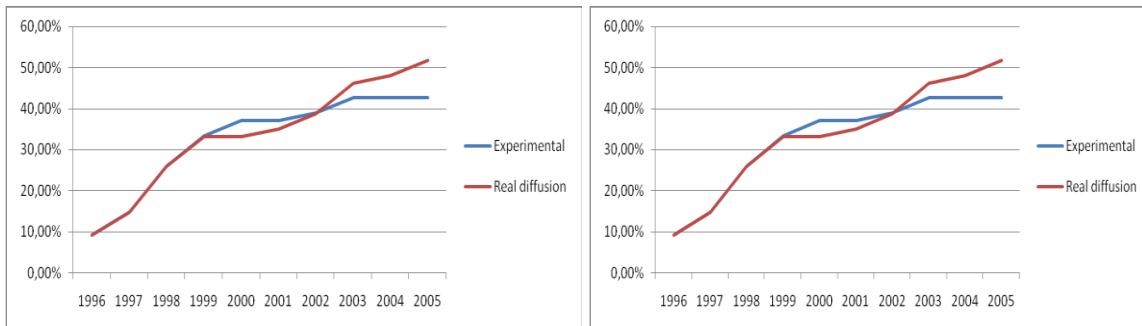


Figure 9: diffusion curves considering 1996-98 history (left) and 1996-99 history (right)

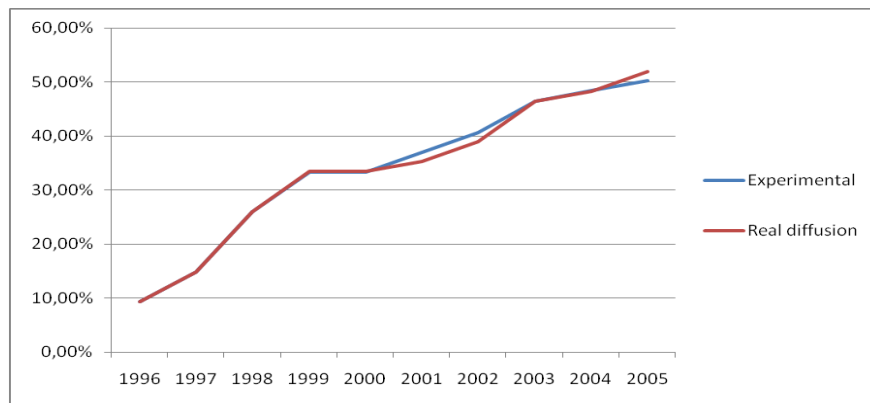


Figure 10: diffusion curves considering 1996-2000 history

Table 3: projection performance evolution

Considered years	96	96-97	96-98	96-99	96-00
% Similarity	63%	72%	81%	89%	92%
Relative Error	12,42%	6,69%	5,38%	4,50%	1,36%

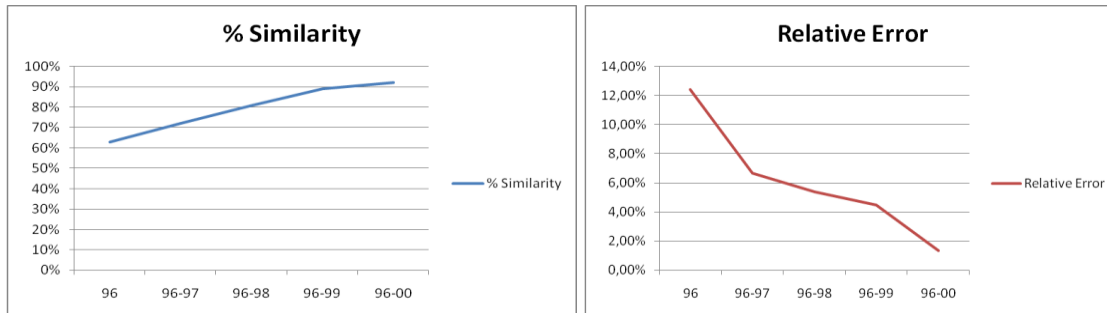


Figure 11: projection performance evolution

## 7. Conclusion

Based on empirical evidence from the adoption of ERP systems by top 54 firms in the automotive sector from 1994 to 2005, we investigated the robustness of a micro diffusion model taking into account the individual perception of the innovation in the North American automotive business network. Based on this real scenario, a simulation-based optimization approach was proposed to tackle the problem. The conducted experimentations have demonstrated the usefulness of such tool for ERP systems vendors to represent faithfully the diffusion phenomenon and to be able to predict the future adopters with a very promising similarity rate.

It is interesting to note that the proposed model and approach could be a good basis for ERP systems editors to test different strategies for market penetration. In this context, being able to target the influential lead users in the initial stages of the diffusion process can be determinant in the success of a new product introduction. This issue is under study and the preliminary results are very promising.

Our next goal will be also to improve our analysis by including additional network properties for the automotive sector. In fact, the proposed model falls to detect random unpredictable market events. Thus, additional parameters in the model should be considered to generalize the findings. As it may interest the reader to know, the same study is under investigation for the North American pharmaceutical business network.

## References

- Abrahamson, E. 1996. "Management fashion," *Academy of management review*), pp 254-285.
- Abrahamson, E., and Rosenkopf, L., (1997). [Social network effects on the extent of innovation diffusion: a computer simulation. \*Organization science\*, 8 \(3\), 289-309.](#)
- Aubert, B., Léger, P.-M., and Larocque, D. 2012. ["Differentiating weak ties and strong ties among external sources of influences for enterprise resource planning \(ERP\) adoption," \*Enterprise Information Systems\* \(6:2\), pp 215-235.](#)

- [Baker, J. E. 1985. Adaptive selection methods for genetic algorithms. Proc. ICGA 1, pp 101-111.](#)
- [Baker, J. E. 1987. Reducing bias and inefficiency in the selection algorithm. Proc. ICGA 1, pp 14-21.](#)
- [Bohlmann, J. D., Calantone, R. J., and Zhao, M., \(2010\). The effects of market network heterogeneity on innovation diffusion: an agent-based modeling approach. J Product innovation management, 27, 741-760.](#)
- [Caruana, R. A., Eshelman, L. A., and Schaffer, J. D., 1989. Representation and hidden bias II : Eliminating defining length bias in genetic search via shuffle crossover. In Eleventh Int. Joint Conf. on AI, Sridharan N. S. \(Ed\), Vol. 1, pp 750-755, Morgan Kaufmann.](#)
- [Chatterjee, R., and Eliashberg, J., \(1990\). The innovation diffusion process in a heterogeneous population: A Micro-modeling approach. Management Science, 36 \(9\), 1057-1079.](#)
- [Chipperfield, A. J., Fleming, P. J. and Fonseca, C. M. \(1994\). «Genetic Algorithm Tools for Control Systems Engineering», Proc. Adaptive Computing in Engineering Design and Control, Plymouth Engineering Design Centre, 21-22 September, pp. 128-133.](#)
- [Choi, T-M., Chow, P-S., and Xiao, T. \(2012\). Electronic price-testing scheme for fashion retailing with information updating. International Journal of Production Economics, \(140:1\), pp. 396-406.](#)
- [Coleman, J. S., Katz, E., & Menzel, H. \(1957\). The Diffusion of an Innovation among Physicians. Sociometry, 20\(4\), 253-270.](#)
- [Davenport, T. H. 1998. "Putting the Enterprise into the Enterprise System," \*Harvard Business Review\* \(76:4\), pp 121-131.](#)
- [Davis, G. F. 1991. "Agents without principles? The spread of the poison pill through the intercorporate network," \*Administrative Science Quarterly\*\), pp. 583-613.](#)
- [Deephouse, D. L. 1996. "Does isomorphism legitimate?," \*Academy of management journal\* \(39:4\), pp. 1024-1039.](#)
- [DeNooy, W., Mrvar, A., & Batagelj, V. \(2005\). Attributes and Relations: Exploratory Social Network Analysis with Pajek.\(Structural Analysis in the Social Sciences\) New York: Cambridge University Press.](#)
- [DiMaggio, Paul J., and Walter W. Powell 1983. "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields," \*American Sociological Review\* \(48:2\), pp. 147-160.](#)
- [Fichman, R. G. 2004. "Real options and IT platform adoption: Implications for theory and practice," \*Information Systems Research\* \(15:2\), pp 132-154.](#)
- [Goldberg, D., 1989. Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, reading. MA.](#)

- [Hajji, A., Pellerin, R., Léger, P.M., Gharbi, A., and Babin, G. \(2012\). "Dynamic Pricing models for ERP systems under network externality", \*International Journal of Production Economics\*, \(135:2\), pp. 708-715.](#)
- Hestermann, C., Pang, C., and Montgomery, N. 2012. "Magic Quadrant for Single-Instance ERP for Product-Centric Midmarket Companies," p. 50.
- [Hwang, Y. 2005. "Investigating enterprise systems adoption: uncertainty avoidance, intrinsic motivation, and the technology acceptance model," \*European Journal of Information Systems\* \(14:2\), pp 150-161.](#)
- [Iyengar, R., Van den Bulte, C., and Valente, T. W. 2011. "Opinion leadership and social contagion in new product diffusion," \*Marketing Science\* \(30:2\), pp 195-212.](#)
- [Kwak, Y. H., Park, J., Chung, B. Y., and Ghosh, S. 2012. "Understanding end-users' acceptance of enterprise resource planning \(ERP\) system in project-based sectors," \*Engineering Management, IEEE Transactions on\* \(59:2\), pp 266-277.](#)
- Léger, P.-M., Babin, G., & Dubois, J.-F. (2008). The Longitudinal Effects of Exposure in Network Diffusion of Enterprise Resource Systems.
- [Li, M., Feng, H., Chen, F., and Kou, J. \(2013\). Numerical investigation on mixed bundling and pricing of information products. \*International Journal of Production Economics\*, \(144:2\), pp. 560-571.](#)
- [Light, & Papazafeiropoulou, A. \(2004\). Reasons Behind ERP Package Adoption: A Diffusion of Innovations Perspective, in Proceedings of the 12th European Conference on Information Systems Turku, Finland.](#)
- [Markus, M. L., Tanis, C., and Fenema, P. C. v. 2000. "Enterprise resource planning: multisite ERP implementations," \*Commun. ACM\* \(43:4\), pp 42-46.](#)
- [Millaire, J.F., Léger, P.M., Babin, G., and Pellerin, R., \(2009\). The effect of network exposure on the diffusion of enterprise system. Proceedings of the Hawaii International Conference on System Sciences, vol. 42. Computer Society Press, IEEE, Hawaii \(2009\) pp. 1–10.](#)
- [Park, S-J and Hahn, M \(2000\). Re-interpretation of the diffusion model: a microeconomic approach. INFORMS-KORMS, Seoul, Korea, 18-21 June, pp. 1065-1070.](#)
- [Pellerin, R., Léger, P.-M., Babin, G. \(2007\), "The impact of board interlocks in the diffusion of Enterprise Resource Planning systems", \*International Journal of Networking and Virtual Organisations\*, Vol. 4, No. 4, pp. 402-412.](#)
- [Pritsker, A. A. B and O'reilly, J. J \(1999\). Simulation with Visual SLAM and Awesim. Wiley, John Wiley et Sons, inc, NY.](#)
- [Shaw-Ching, L. B., Ravindranath, M., & Sudharshan, D. \(2005\). DiffuNet: The Impact of Network Structure on Diffusion of Innovation. \*European Journal of Innovation Management\*, 8\(2\), 240-240.](#)
- Strang, D., and Soule, S. A. 1998. "Diffusion in organizations and social movements: From hybrid corn to poison pills," *Annual review of sociology*, pp 265-290.
- [Syswerda, G., 1989. Uniform crossover in genetic algorithms. Proc. ICGA 3, pp 2-9.](#)

- Trunda, T., & Westarp, F. (1998). Edi Solutions in Large Companies-a Case Study with 3com.
- Valente, T. W. (1995). Network Models of the Diffusion of Innovations. Cresskill: NJ: Hampton Press.
- Van Eck, P.S., Jager, W., and Leeflang, P.S.H. (2011). Opinion leaders' Role in innovation diffusion: a simulation study. J. Product innovation management. 28, 187-203.
- Wedel, M., and Kamakura, W.A. (2000). Market segmentation: Conceptual and Methodological Formulations. (2d ed.). Boston: Kluwer Academic Publishers.
- Wybo, M., Robert, J., & Léger, P. M. (2005). An Optimization Model of the Business Applications Selection Process. Cahier du GReSI no, 5, 08-08.