Exploring Network Modelling and Strategy in the Dutch Product Software Ecosystem

Wesley Crooymans[1], Priyanka Pradhan[1] and Slinger Jansen[2]

Utrecht University, Buys Ballot Laboratorium, Princetonplein 5, Utrecht, Netherlands
1 {W.A.S.Crooijmans, P.Pradhan}@students.uu.nl
2 slinger.jansen@uu.nl

Abstract. In today's product software market, the practices of re-use, partnering and 3rd party contracting give rise to complex software ecosystems. Over the duration of a product life-cycle, product software vendors build up relationships with their suppliers and other partners, which range from informal acknowledgements of each other's presence to strategic alliances. There is still a lack of understanding surrounding the roles, connections, relationships, and resulting networks within software ecosystems. Using modelling techniques and statistical analysis, these networks can be used as tools to further that understanding. In this paper a collection of 67 software supply networks will be modelled as a network graph. Using clustering and two extensions of basic software supply network data, we identify several major players and domains in the Dutch software industry. Three business strategy perspectives are then related to the data to provide an example of their potential practical use.

Keywords: Software ecosystems modelling; network modelling; product software; software business strategy

1 Introduction

Problem Statement. A Software Ecosystem (SECO) model can be a powerful aid to describe the position and role of a software business within its environment[1]. With various modelling methods and algorithms available today, there is no shortage on fancy visuals and beautiful graphics to shed a light on your ecosystem. The question to ask is, *How can software ecosystem models be used to benefit businesses?* In this paper we will answer that question by presenting a method for modelling, analyzing and interpreting ecosystem data, applied to a study of the Dutch product software industry. To portray some of the ways in which these models may be of use as tools to the business world, the data will be related to existing papers on (SECO) business strategy.

Domain. The work in this paper lies firmly within the domain of SECOs. Throughout this paper the definition of Jansen, Finkelstein, and Brinkkemper (2009) will be used: "We define a SECO as a set of businesses functioning as a unit and interacting with a

adfa, p. 1, 2011.

shared market for software and services, together with the relationships among them. These relationships are frequently underpinned by a common technological platform or market and operate through the exchange of information, resources and artifacts.”[2]

In a SECO there are number of strategies an organization can adopt. The kind of strategic decisions that are available depend on the individual organization and their perspective on the ecosystem. Iansiti and Levien stated that there are three roles that influence ecosystem health and evolution. Some known strategies are partnering, membership programs, mergers and acquisition etc. For example Cisco used merger and acquisition strategy to be the keystone player in their ecosystem, which helped them to foster success for the whole ecosystem as well as to generate revenue [3].

Our area of interest lies specifically within the product software area of SECOs, and the relationships among the various suppliers, vendors, retailers and customers surrounding software products. Software products in this paper include all software packages and services that are traded as standard products [4], including but not limited to Consumer-Of-The-Shelf (COTS), standard software, shrink-wrapped software, Software as a Service (SaaS), open source software, and any packages or configurations of the former.

One way of modelling these SECO relationships, is the so called Software Supply Network (SSN) model. A software supply network, according to Jansen, Brinkkemper and Finkelstein (2007), is “a series of linked software, hardware, and service organisations cooperating to satisfy market demands”[5]. The accompanying modelling method describes SSNs in two parts; the product context and the supply network. The product context describes which products and services make a complete software product, and a supply network describes all parties involved in the SSN including the flows of goods, services and finances between them [5]. The supply network notation from this modelling method was also used to create the dataset used throughout this paper, which will be described in more detail in the method section.

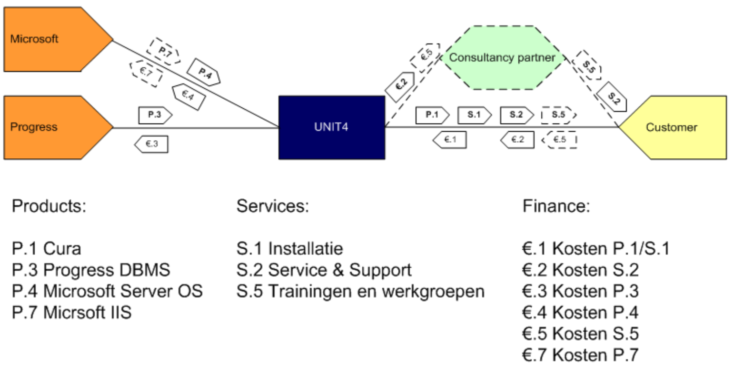

Fig. 1. SSN Supply network example

Suppliers of software, hardware, services or content are drawn as orange rectangles with a pointed right edge. Intermediaries that act as partners, implementers or resellers are drawn as green rectangles with two pointed edges. Customers are represented as yellow rectangles with a pointed left edge. Finally, the vendors of a product are the central focus object of an SSN, and these are drawn as blue rectangles. Flows between elements in an SSN are represented as simple lines, annotated with the contents of the flow in pointed rectangles. See figure 1 above for an example.

Outline. Section 2 covers the clean-up and preparing of data for the rest of the paper, as well as the methods used for analysis. In section 3 the dataset will be described in more detail. This section will also cover the process of modelling the data in Gephi and the methods used to visualize the data. The additions made for the 2013 extended case and Microsoft case are also be described. Section 4 will present the results of community detection and preliminary findings from the extended 2013 case and Microsoft case. In section 5 these findings will be related to SECO strategy literature in academia. The paper concludes with a discussion in section 6 and a conclusion in section 7.

2 Method

In this section, we elaborate on the methods used in this paper, like the collection of data, the process of conducting statistical analysis and the application of analytical methods.

Data. The data source for this research was obtained from several SSN modelling case studies performed in the product software market in the Netherlands. These case studies were conducted by Utrecht University undergraduate students as part of a product software course. Our data set consists of supply network models created over three years (2010, '11 and '13).

As a part of the undergrad course, students were asked to perform interviews at a company selling software products. The only criterion on company choices was ownership in the Netherlands. The purpose of these interviews was to investigate how companies built and sold their products, to the point that SSNs could be created by the students. Not all students performed their assignments up to the same standards of quality. To assure that only high quality data was available for this project, several criteria were used to exclude certain low quality cases:

1. **Unnamed partners**. In cases where a lot of suppliers and/or intermediaries were not given a proper company name (e.g. 'hardware vendor', 'consultant'), the data was considered unusable and the case was removed.
2. **Illogical flows.** In some cases students failed to accurately model the flow of products, services and finances between parties in the SSN (e.g. finances flowing from A to B for no product or service from B to A). Singular illogical flows were removed from the cases. Large numbers of illogical flows led to case removal.
3. **Irrelevant for the vendor ecosystem**. In some cases students included suppliers that had no interaction with the product or the vendor providing it (e.g. A supplier selling servers to the customer directly). Suppliers that had no meaningful contribution to the product or relationship with the vendor were removed from their case.

In the 2013 iteration of the course, students were also tasked with registering several attributes about the vendors' relationships with other companies in a central datasheet. Four of these attributes were of particular interest to us: The relationship type with a party, the perceived balance of power in that relationship, the perceived importance of that relationship to the vendor's business model, and the frequency of contact with the other party.

The basic data for exploring the Microsoft ecosystem were obtained from the same source. The original data were filtered so that only the 101 parties with a connection to Microsoft remained. 50 Organizations were listed in the Microsoft PinPoint search engine. 47 organizations either claimed a relationship on their website, or made no information available. Only, 10 out of these 47 organizations replied to an e-mail with the information needed. Another two organizations were not available for contact and two more were acquired by Microsoft since the SSNs were made.

Visualization. After collecting data, the open source tool Gephi is used for visualization. Gephi has a flexible and multi-task architecture that allows filtering, navigating, manipulating and clustering for complex graphs of network models [6].

The SSN supply networks from the case studies were converted and used as input parameters. Parties and relationships from the models are entered as the nodes and edges of a network. For the example supply network in figure 1, the nodes and edges in the resulting data tables are visualized in table 1 and 2 respectively.

Table 1. Node table structure

Node	ID	Label	Type	SSN
Microsoft	1	Microsoft	Supplier	1
AIP	2	AIP	Vendor	1
AIP_Customer	3	AIP_Customer	Customer	1

Each node in Table 1 is given a name (Node) and a unique ID. The Type attribute denotes the role of a node in the SECO. The SSN number makes it possible to identify which SSN every node in the network originated from.

Table 2. Edge table structure

Source	Target	Type	ID	Weight	Products	Services	Finances	Content	Hardware
1	2	Directed	1	5	5				
2	1	Directed	2	3			3		
2	3	Directed	3	2	1	1			

Edges carry the Source- and Target-IDs of the nodes that they connect. Gephi automatically gives these source-to-target relationships the Type attribute 'Directed'. The total amount of resources exchanged within a single relationship (i.e. products, services, finances, content and hardware) forms the Weight of the relationship.

For the 2013 dataset the edge tables were expanded with the four relationship attributes (partner importance, relationship type, contact, and balance of power), converted to integer values from their original Likert scale rankings. A partner with 'monthly' contact would, for instance, get a score of 4 on the matching attribute.

Table 3. 2013 extended case relationship attribute coding

Partner Importance		**Relationship type**		**Contact**		**Balance of Power**	
Crucial	5	Cooperation agreement	5	Weekly	5	Partner more powerful	5
Very important	4	Partnership program	4	Monthly	4		
Important	3	License agreement	3	Yearly	3	Power is equal	3
Not very important	2	Informal relationship	2	Rarely	2		
Trivial	1	Animosity	1	Never	1	We are more powerful	1

Edge weights were recalculated as the sum of all four relationship attributes. Edge types were set to 'Undirected', as the direction of edges no longer holds any meaning in this case.

Microsoft certainly seems to be the keystone player of the Dutch ecosystem from previous visualizations. Therefore, we expanded the data related to Microsoft to get the inside view of its ecosystem, as outlined in section 2.1. For Microsoft partnership data structure, the format of nodes is depicted in table 4. The 'partner' attribute describes which type of partnership the organizations has with Microsoft. Partner total is the sum of gold and silver certificates the organization has with Microsoft.

Table 4. Node table structure for the Microsoft case

Node	**ID**	**Label**	**Type**	**SSN**	**Partner**	**Partner Total**
HP	3	HP	Supplier	1	Gold	14
AIP	19	AIP	Vendor	2	Silver	1
Centric	33	Centric	Vendor	3	Gold	11

Edge type is once again undirected, the reason being the same as for the 2013 case. The edge tables were extended with attributes for the amount of gold and silver certifications, and the Weight attribute was made the sum of all certifications.

For the sake of exploration three distinctly different structuring algorithms will be used to draw network graphs throughout this paper. The Fruchterman-Reingold-, Force Atlas-, and Force Atlas 2 - algorithms each offer a different degree of user control, but more control always comes at the cost of ease-of-use.

Clustering. For extracting clusters from the data, the Louvain clustering algorithm is used [7]. The Louvain algorithm is proposed by Blondel, Guillaume, Lambiotte and Lefebvre. This algorithm works by calculating and optimizing the modularity (the density of edges) of groups of nodes. On the first pass of the algorithm, each node is placed in its own community. From there the algorithm follows this pattern:

1. Calculate the modularity of all communities in the network.
2. For each community, for each neighbor, calculate the gain in modularity if it were to merge with that neighbor

3. For the highest detected modularity gain, merge the two neighbors
4. Recalculate modularity for all communities that remain.
5. Repeat from step 1 until no modularity gains are possible.

3 Describing the dataset

After elimination in accordance to the criteria outlined in section 2, a total of seven cases were removed from the dataset for their lacking quality. A total of sixty seven cases remained in the final dataset. After entering the data into Gephi, as outlined in section 2, a network model was created containing 398 nodes connected by 984 edges. Using the Fruchterman-Reingold algorithm for force directed graph drawing[8] resulted in the network shown in figure 2. Fruchterman-Reingold forces nodes to be placed within a predefined circular space, with only minor variations in inter-node distances based on the weight of their connections. This makes the resulting graphs easy to create and control, particularly for large networks. The same algorithm was used by Rahul Basole in his visualizations of the mobile ecosystem [9]

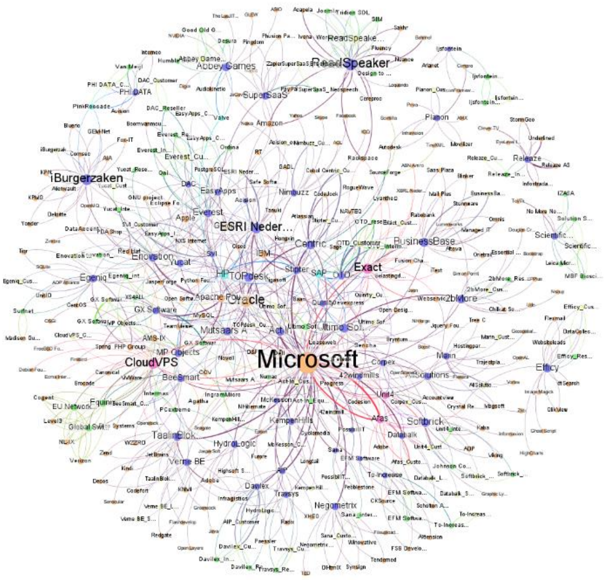

Fig. 2. Full network diagram as generated by combining all the data

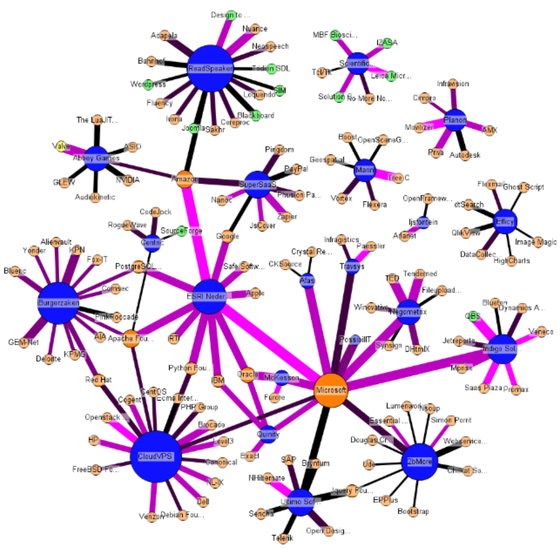

Fig. 3. 2013 extended case diagram with edge brightness reflecting contact frequencies

This diagram was created using the following parameters: Node size = degree (amount of incoming and outgoing connections); Edge size = Edge weight

The network has the following composition: 222 suppliers (orange), 52 vendors (blue), 56 intermediaries (green), 60 customers (yellow), 4 supplier/vendor hybrids (purple), and 4 supplier/intermediary hybrids (teal).

A total of 572 products change hands in the network. In addition, 191 services are provided and 6 units of hardware are sold. In return, 40 pieces of content are provided and 521 payments are made. In total the network captures 1329 exchanges being made.

Perhaps unsurprisingly, the best connected node in the entire network is Microsoft, with a total of 79 unique connections. Microsoft is however not the only strongly connected supplier in the dataset. What follows is a top 10 list of the best connected suppliers:

1. Microsoft	6. SAP
2. Oracle	7. IBM
3. Apache Foundation	8. Apple
4. HP	9. AMS IX
5. Google	10. Amazon

The top 3 of the Forbes Global 2000 for software companies in 2013 are all present, namely Microsoft, Oracle and SAP. Also the number one and two hardware providers, (Apple and HP), and the number one and two software service providers, (IBM and Google), made the list[1]. Finding these giants at the centre of ecosystems is not uncommon in ecosystems research [10, 11] Perhaps a good sign of health for the open source software market is the Apache Foundation ranking 3rd. Other open source providers include but are not limited to the Eclipse foundation, The Python foundation, Debian, FreeBSD, Red Hat and the GNU project. The only natively Dutch supplier to make the list is the Amsterdam Internet Exchange (AMS IX).

In the extended 2013 case 152 nodes and 155 edges remained. Using Gephi's own Force Atlas 2 structuring algorithm results in a structure that makes it easier to identify outliers, compared to Fruchterman-Reingold [12]. Force Atlas 2 uses attraction, repulsion and scaling variables to iteratively determine structure. Force Atlas 2 and its predecessor were similarly used by Yu, Yin, Wang and Wang to visualize social groups in the Github ecosystem [13] Default values will often result in groups of connected nodes clustering too tightly and outliers flying off into the void, so some tweaking on part of the user is required. The resulting graph can be seen above in figure 3. This diagram was created using the following parameters: Node size = Degree; Node color = Same as Figure 2 (Valve is yellow); Edge size = Edge weight; Edge color = 5 point gradient for Contact frequency. For this diagram the edges were colored on a 5 point gradient, reflecting the values of the 'contact frequency' attribute. Frequent contact results in a bright purple edge color, whereas no contact whatsoever results in a black edge. Four more diagrams were created with different color schemes for the 'relationship', 'power' and 'importance' and overall 'weight' attributes.

For the Microsoft case, after filtering 55 nodes remained and 54 edges between them. The Force Atlas algorithm was used to structure the data in Gephi [6]. Force Atlas operates on the same principles as Force Atlas 2, but allows users more control. This makes it more flexible for creating visualizations, but also more difficult to use for the inexperienced. The diagram was created using the following parameters: Node size =

[1] http://goo.gl/tir5Tx

node weight; Edge size = Edge weight. Nodes for partners with gold and silver partnerships have colors to match. There are also two acquisitions by Microsoft (green) and four partners without certifications (red). There are 48 partners with Gold and/or Silver certificates, 4 partners without certificates and 2 companies are acquired as can be seen in Figure 4.

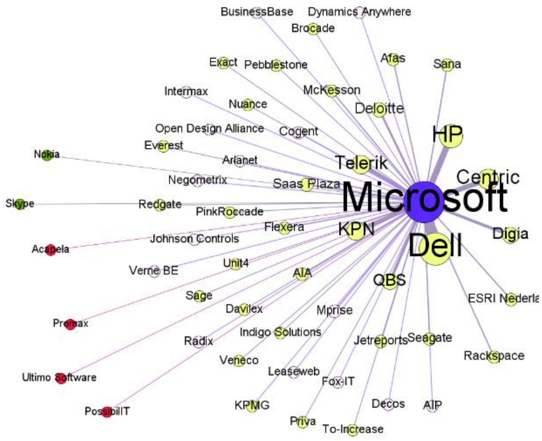

Fig 4: Full Microsoft network diagram

Microsoft has 33 Gold certified partners in our dataset. The organization with the most certifications is Dell with a total of 25 (21 gold) certificates. There are 15 partners in the network with *only* silver certificates. Comparatively there are more gold certified partners than silver ones. The possible reason for this will be discussed in section 6.3.

4 Preliminary findings

In this section the preliminary findings of modelling the dataset are discussed. It will discuss the results of clustering, extending the data for 2013, and extending the data for Microsoft respectively.

Clustering. After running the Louvain Algorithm 21 communities were detected within the full dataset. 10 of these are identical to their original SSN diagrams, as described in section 2. Vendors that are only tied to the larger network through single edges can achieve the greatest modularity gain by simply merging into a single node. Five of the detected communities consist of two merged SSNs. This pattern occurs when two central vendors in the SSNs share their strongest connection with a single supplier or intermediary. When this occurs the Louvain detects the highest possible modularity gain in merging those SSNs. This leaves 6 more complex communities:

The Microsoft Community - The largest,is a community centered on Microsoft. Software vendors that do not have ties with any other large suppliers tend to get sucked into this community because it is the only way for the algorithm to improve their modularity.

A taste of Open Source (fig 5) - The second largest community is all about open source. It appears that vendors with ties to one open source supplier tend to get a good taste for more. The central nodes in the community include Red Hat, the Apache Foundation and the Eclipse foundation. Other closed source suppliers also get pulled into this community because of their strong ties with the open-source-using vendors.

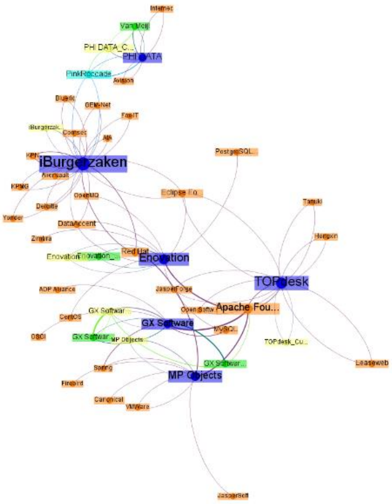

Fig. 5. The open source community

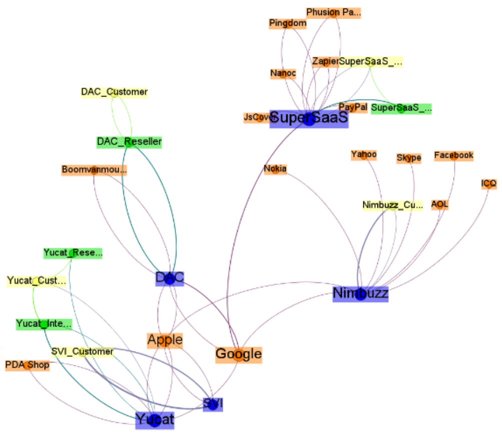

Fig. 6. The mobile developers community

The Mobile Developers (fig 6) – The third largest community centers on Google and Apple. Digging back into the original SSNs reveals what drives this community. iOS and Android. Those vendors that consider mobile platforms a crucial component of their products are pulled into this community.

The Cloud Providers – The fourth largest community is up in the clouds, whether its hosting services, data centers or specific implementations like video-on-demand. This community centers around CloudVPS, Equinix and the PHP group, providing servers, datacenters and a server-side scripting language respectively.

The Oracle community (fig 7) – The fifth largest community centers on Oracle and includes Autodesk. This may be in part due to a strategic partnership established between the two suppliers in early 2007. Their single shared connection to Centric however, does not provide any conclusive evidence for this assumption. Other vendors that are well connected to Oracle also get sucked in.

The Hardware community (fig 8) – The sixth and last complex community centers on established hardware vendors. Three big names immediately stand out: HP, IBM and Cisco. In an age of increasingly affordable hardware, SaaS and Cloud solutions, some vendors still rely on services of these big name suppliers.

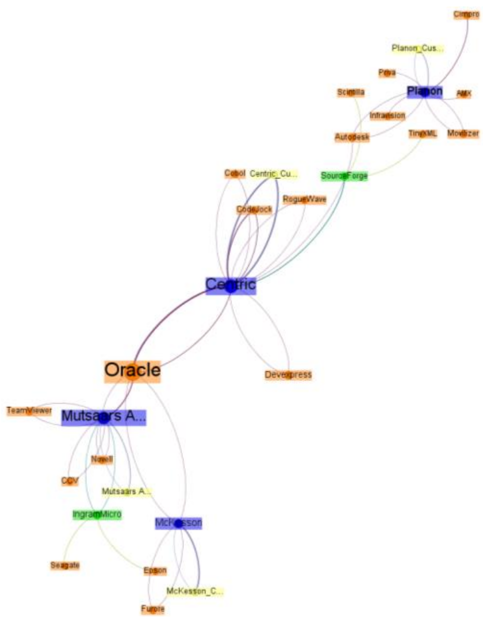

Fig. 7. The Oracle community

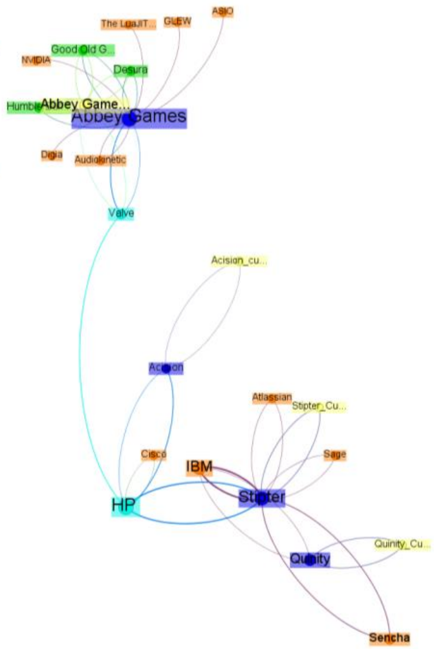

Fig. 8. The Hardware community

2013 case. An immediate pattern that arises from the 2013 case models (see figure 3), is the large discrepancy between perceived partner importance for the product and contact frequencies/relationship types with that partner. One marketing software vendor for example, has indicated that Microsoft is absolutely crucial to their product (5/5 rating). When asked on their relationship with Microsoft however, the vendor has indicated that there is none (1/5 rating) and they rarely have contact with Microsoft (2/5 rating). This is in an unexpected contrast, and is repeated on several occasions by others within the network.

On the 1 to 5 rating scale, the differences between importance and relationship ratings (n=155) averaged -.08 (s=1.946). There are 15 outliers (rating difference =>3) for which, despite the high importance of a partner, the relationship with them is informal. The differences between importance and contact frequency ratings (n=155) averaged -.11 (s=1.689). The amount of outliers (rating difference => 3) in this category is exactly the same.

These differences may be indicative of a strategic gap. It is unexpected for software producing companies to have little to no contact and informal relationships with their most important suppliers. There may however, be other factors that can explain these differences. Bigger suppliers may be less available for contact and partnership programs may be expensive to enroll in. Particularly for smaller companies.

Looking back for a moment at the original SSN models, there were no indicators that these differences even existed. In the SSN notation all relationships between ecosystem parties are equal. The statistics above show that relationships between parties cannot be assumed to be the same. Whether this indicates a need for change in the SSN notation, or whether this change is outside of the intended scope, is a question for future research. An attempt to include this sort of data in SSN models has been made before by Handoyo, et al. [14].

Microsoft case. As can be seen in figure 4, the numbers of gold partners are comparatively higher than silver partners. There may be few reasons for the organizations to become gold certified rather than silver. One of the reasons may be differences in benefits making gold certification more attractive. For instance, Microsoft gives partners free licenses for the internal use of some Microsoft products; silver partners can use only up to 25 licenses per product while gold partner can use up to 100 licenses per product. Also, the enrollment procedures for Gold and Silver certificates are not that different. In both cases employees of the prospective partner must take exams and pay license fees (though higher for gold). A minor difference for example: Gold certified organizations must use the customer satisfaction (CSAT) index survey for performance measurement, which takes extra work but has its own benefits This may contribute to an attitude that one looking to get certified might as well just ‘go for Gold’.

5 Relating the data to business strategy

It is worth noting that even without analysis the dataset and accompanying models already provide interesting information. The model shown in fig. 2 shows SECO structure, basic information on the SECO parties, their roles, and relationships between them. Much of SECO strategy literature in academia already uses this data to identify, classify, create and assess strategies in the business world.

When SECO roles are concerned, most literature will name keystones [15, 16] and niche players [17, 18] as the two most prominent. Both of which can be easily identified in a network diagram by looking at node centrality and modularity. The role of a company has significant influence on their strategies with regard to their ecosystem.

Ecosystem structure also plays a role in SECO strategy literature. Van den Berk, Jansen and Luinenburg name it as a factor in a SECO strategy assessment model [1], and Iyer, Lee and Venkatraman use structural data as a measure for SECO health [19].

Relationships in ecosystems form the basis for research on partnership programs. For instance Bosch who defines strategic decisions for ecosystem partners to make the most out of their (potential) relationships [20]. Another example is given by Popp, who identifies goals related to partnership programs and communities [21].

Communities. In their 2011 paper on SECO management practices, Viljainen and Kauppinen synthesized four major categories of practices [22]. Three of these categories can play a role in business strategy with model support. Some of those practices are the following:

Technology Scouting - The communities detected in section 4 can be used to support these practices. Having an understanding of one's local community, including its structure and other participants, can provide a basis for choosing targets in technology scouting practices. Particularly targets for joint ventures and acquisitions can be justified by evidence of a shared community.

Orchestration - For keystone players interested in orchestration, the communities can serve as inspiration for identifying closely related parties that they were not previously aware of, potentially extending the boundaries of their perceived SECO. These parties can then be targeted for partnerships and standards adoption.

Technology asset management - For those looking to change up their technology asset management practices, communities can help to identify parties with valuable knowledge and similar practices. Vendors in the open source community could for instance find other vendors that use the same open source components. Sharing their requirements and knowledge can help both parties to get the most out of their open source components.

2013 case. The 2013 extended case provides interesting information on vendor perceived metrics of ecosystems. The data includes partner importance, relationship type, contact frequency, and the balance of power. These metrics could be of particular interest to research in ecosystems health.

Hartigh, Tol and Visscher set out to create formal measures for the concept of SECO health [23]. Their work extended on an earlier paper by Iansiti and Levien, who already defined productivity, robustness and niche creation as the three main categories of SECO health factors [15]. Under the factors "persistence of ecosystem structure" and "predictability" in the ecosystem Robustness category, Hartigh et al. name measures like the amount of connections of each agent and the 'connectedness' of the entire network as measures.

In section 4 large discrepancies between relationship importance and relationship types were detected. This begs the question whether a simple tally of relationships and a number for connectedness are sufficient to measure robustness. The use of such values can certainly be valuable indicators of SECO health, but they do not paint a complete picture. Twenty or more connections are hardly an indicator of network robustness if those connected perceive their relationship as informal. This shortcoming in their measures was also noted by Hartigh et al. Using data and models similar to the 2013 case could help to further improve the measurement of SECO health.

Microsoft case. For the Microsoft case the partner ecosystem defined by Popp is particularly interesting [21]. In an earlier work Meyer identified the following categories of goals for a software vendor in a partner ecosystem: financial, customer related, product related, network effect related and market related goals [24]. Popp makes these goals explicit for his view of the partner ecosystem. Microsoft already takes its partner ecosystem goals very seriously, as is said on their website: “Microsoft believes that their own advantage shines through the success of their partners” [25]. Microsoft’s mission includes partners as a central means to help customers and business throughout the world.

Under product related goals innovation and co-innovation in local and regional markets could be helped by visualizing networks that include geographical data. This does not just show partners on a map, but also makes it easier to identify local and regional keystones and niche players. This can be used to analyze their strongest potential partners for co-innovation in a certain region, and to identify the best partners for reselling Microsoft products in a region.

In the category of financial goals Microsoft gains more from gold certified partners than from silver certified partners. Accurate information about silver certificates in a market as shown in figure 4, could help Microsoft find partners for upgrading to a gold certification in the same competency. These partners could then be contacted or even helped to achieve an upgrade.

6 Discussion

In this paper we discussed a dataset, several modelling approaches, several analysis results and several ways of relating these results to SECO strategy goals and practices from literature. Though there is some research available on SECO modelling and the use of those models, there is as of yet very little to be found on the inter-seco level of ecosystems [2]. On the inter-seco level, ecosystem models provide a view of the connections between multiple vendors, suppliers and intermediaries, rather than taking a single vendor as the Company of Interest as the SSN notation does.

The major validity concern in this paper is the source of data that was used. Three years of bachelor’s course results are unlikely to all be of equal quality and reliability. Some filtering was done to exclude particularly bad data, as outlined in section 2, but there may still be faults left. The validity of the data depends on the assumption that all remaining students were honest, objective and diligent in their work.

Regardless, the specific contents of the dataset do not directly threaten the validity of the methods used in sections 4 and 5. Even with a perfect dataset the results would likely still have indicated gaps between relationship importance and contact/strength, community clusters of a similar nature, and similar Microsoft Partnership program results. For future work a systematic data collection method can be used to create a new dataset.

This paper succeeded in showing the potential use of ecosystem models for the business strategy field. There are many other focus areas in the SECO field where models may also contribute to existing practices and goals. A sample of other fields for future research is provided in the conclusion of the paper.

The uses of models found can help businesses to create more robust and complete strategies and decisions. There may also be some benefits to be had for other parties.

Views of SECOs could for instance help policy makers and market researchers in understanding the product software market as a whole. This is a good subject for future research.

7 Conclusion

The original research question was '*how can software ecosystem models be used to benefit businesses?'*. The first step towards answering this question was the creation of models. SSNs gathered from 3 years of bachelor level courses in product software were used to create a view of the Dutch product software ecosystem. This view was then analyzed using a community detection algorithm, and extended using more specific data about relationships and partnerships.

To show that the three created models and the analysis of those models can actually have a benefit for business, the results section tied the data to strategic practices and goals described in scientific literature. The community detection results were related to three categories of SECO management practices, the 2013 case was related to a SECO health measurement framework, and the Microsoft partners case was related to goals for successful SECO partnership models.

In this paper the strategic perspective was kept fairly general. For future research similar work can be done to show the potential uses of models for focus areas within the SECO strategy domain. Some examples of focus areas are ecosystem governance and orchestration, SECO health analysis, partnership management, and software platform development. These subjects were touched upon briefly from the strategy perspective in this paper, but could be further expanded upon in later works. Another option is to move away from software businesses entirely, and to investigate the use of these models for external parties (e.g. governments, market research, business analytics).

Another opportunity for future research that was identified in this paper is the use of relationship attributes in SSN modelling and SECO health measurement. The potential shortcomings of SSNs without relationship information were discussed in section 4. The lack of relationship data in health measurement was discussed in section 5. Both fields could benefit from the realization that not all relationships are equal when considering the formality and contact frequency of those relationships.

References

1. Van den Berk, I., Jansen, S., Luinenburg, L.: Software Ecosystems: a Software Ecosystem Strategy Assessment Model. Proc. of the 4th ECSA: Companion Volume. pp. 127–134. ACM, New York, NY, USA (2010).
2. Jansen, S., Finkelstein, A., Brinkkemper, S.: A sense of community: A research agenda for software ecosystems. Proc of the 31st ICSE: Companion Volume. pp. 187–190 (2009).
3. Li, Y.-R.: The technological roadmap of Cisco's business ecosystem. Technovation. 29, 379–386 (2009).
4. Xu, L., Brinkkemper, S.: Concepts of product software. Eur. J. Inf. Syst. 16, 531–541 (2007).

5. Jansen, S., Brinkkemper, S.: Providing Transparency In The Business Of Software: A Modeling Technique For Software Supply Networks. Proc. of the 8th IFIP Working Conference on Virtual Enterprises. pp. 667–686 (2007).
6. Bastian, M., Heymann, S., Jacomy, M., others: Gephi: an open source software for exploring and manipulating networks. 3rd AAAI CWSM. pp. 361–362 (2009).
7. Isaksen, A.: Knowledge-based clusters and urban location: the clustering of software consultancy in Oslo. Urban Stud. 41, 1157–1174 (2004).
8. Fruchterman, T.M.J., Reingold, E.M.: Graph drawing by force-directed placement. Softw. Pract. Exp. 21, 1129–1164 (1991).
9. Basole, R. C. (2009). Visualization of interfirm relations in a converging mobile ecosystem. Journal of Information Technology, 24(2), 144-159.
10. Iyer, B., Lee, C. H., & Venkatraman, N. (2006). Managing in a small world ecosystem: Some lessons from the software sector. California Management Review, 48(3), 28-47.
11. Iansiti, M., & Levien, R. (2004). Keystones and dominators: Framing operating and technology strategy in a business ecosystem. Harvard Business School, Boston.
12. Jacomy, M., Venturini, T., Heymann, S., Bastian, M.: Forceatlas2: A Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. PLoS One. 9, e98679 (2014).
13. Yu, Y., Yin, G., Wang, H., & Wang, T. (2014, November). Exploring the patterns of social behavior in GitHub. In Proceedings of the 1st In-ternational Workshop on Crowd-based Software Development Methods and Technologies (pp. 31-36). ACM.
14. Handoyo, E., Jansen, S., Brinkkemper, S.: Software ecosystem modeling: the value chains. Proc. of the 5th ICMEDES. pp. 17–24 (2013).
15. Iansiti, M., Levien, R.: The keystone advantage: what the new dynamics of business ecosystems mean for strategy, innovation, and sustainability. Harvard Business Press, Boston, MA (2004).
16. Riis, P.H., Schubert, P.: Upgrading to a new version of an erp system: a multilevel analysis of influencing factors in a software ecosystem. Proc. of the 45th HICSS. pp. 4709–4718 (2012).
17. Kazman, R., Chen, H.M.: The metropolis model: a new logic for development of crowdsourced systems. Commun. ACM. 52, 76–84 (2009).
18. Kabbedijk, J., Jansen, S.: Steering insight: An exploration of the ruby software ecosystem. Software Business. pp. 44–55. Springer (2011).
19. Iyer, B., Lee, C.H., Venkatraman, N.: Managing in a small world ecosystem: Some lessons from the software sector. Calif. Mgmt. Rev. 48, 28–47 (2006).
20. Bosch, J.: From software product lines to software ecosystems. Proc. 13th Int. Softw. Prod. Line Conf. 111–119 (2009).
21. Popp, K.M.: Goals of Software Vendors for Partner Ecosystems--A Practitioner's View. Software Business. pp. 181–186. Springer (2010).
22. Viljainen, M., Kauppinen, M.: Software ecosystems: A set of management practices for platform integrators in the telecom industry. Lect. Notes Bus. Inf. Process. 80, 32–43 (2011).
23. Den Hartigh, E., Tol, M., Visscher, W.: The health measurement of a business ecosystem. Proc. of the ECCON 2006 Annual Meeting. pp. 1–39 (2006).
24. Meyer, R.: Business Models for Software Companies. BoD, Hamburg, Germany (2008).
25. Popp, K., Meyer, R.: Profit from Software Ecosystems: Business Models, Ecosystems and Partnerships in the Software Industry. BoD, Hamburg, Germany (2010).