

Mapping network development of international new ventures with the use of company e-mails

Ingrid Wakkee

Published online: 18 May 2007
© Springer Science + Business Media, LLC 2007

Abstract International new ventures use e-mail frequently to communicate with globally dispersed contacts. In this paper we present and discuss a qualitative research method to map international network development based on company e-mails. Our approach also allows for combinations of inductive and deductive analysis and combines the possibility of content analysis and descriptive statistics with sensemaking through qualitative analysis of the narratives unfolding in the e-mail chains between the firm and its alters. We discuss these issues and examine to what extent and how structural and relational network variables and dimensions can be examined using e-mails and why this is relevant. In this discussion e-mails are also triangulated with and compared to other sources of data. The data used to exemplify our method are taken from an in-depth case study of a Dutch International New Venture.

Keywords Qualitative methods · E-mails · Data sources · Social networks · International entrepreneurship · Longitudinal

Introduction

About half of the empirical studies in international entrepreneurship have qualitative designs (e.g., Andersson and Wictor 2003; Bell et al. 2001; Coviello and Munro 1995; McDougall et al. 1994; Oviatt and McDougall 1995; Rasmussen et al. 2001; Sharma and Blomstermo 2003). Most commonly, these enquiries adopt case study approaches building on interview data and secondary sources (Coviello and Munro 1995; Rialp et al. 2005). The analysis is usually based on pattern-matching and explanation-building and sometimes on qualitative profiling, event, network, and critical-incident analysis. These methods seem appropriate for capturing complex phenomena (Coviello and Jones 2004: 495) but “data... are presented in an exploratory

I. Wakkee (✉)
Vrije Universiteit, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
e-mail: iam.wakkee@fsw.vu.nl

and highly narrative manner, thus rarely supported with other figures and/or statistics” (Rialp et al. 2005: 157). Consequently, international entrepreneurship scholars are now calling for more innovative qualitative designs that integrate positivist with interpretivist methodologies and incorporate time as a key dimension (e.g., Bell and Loane 2002; Coviello and Jones 2004; Neergaard and Ulhoi 2006)

Concerning the actual topics of the research on international entrepreneurship, cooperation and social networking have become dominant themes in international entrepreneurship research (e.g., Coviello and Munro 1995; Zucchella 2001; Blomstermo and Sharma 2002; Andersson and Victor 2003; Jones and Dimitratos 2004; Dimitratos and Jones 2005). Typically, social network models suggest that networking facilitates access to knowledge and opportunities, and allows startups to overcome resource constraints and to manage a global reach from inception, without bearing the costs of a global proprietary structure (Zucchella 2001; Coviello and Munro 1995).

Considering the dominance of network approaches, it is important that the new methodologies for analysis are more dynamic in the sense that they incorporate time and context (Easton 1995; Halinen and Törnroos 2005), and that they allow for better combinations of structural and relational analysis than current methods (Podolny and Baron 1997). From our perspective, using company e-mails allows researchers to overcome some of the limitations involved in many other qualitative methods that have been used in studying both new venture internationalization and social networks.

E-mails have been used in previous studies on network development inside organizations (e.g., Kling 1996; Haythornthwaite and Wellman 1998; Krebs 2003) or on personal network development (e.g., Donath 2004). Typically, in these studies the headers (“to”, “from”, “cc”, “bcc”, “subject” and “date” fields) are used to (graphically) represent the structure of developing social networks. Such an approach makes good sense. Several good software packages are available that allow you to model such network structure relatively easily and accurately (Krebs 2003; Donath 2004). Also, as argued by Donath, hidden among its sent, receive, ccs, and bccs is a depiction of one’s social network, of changes that demarcate moves to new jobs and new cities, and varying linguistic features and exchange rhythms that create a portrait of one’s acquaintances and their relationships. In other studies, the content of the messages has been investigated as well, using a variety of methodological approaches. A nice overview of the approaches is provided by Donath (2004). She includes:

- semantic analysis, which is based on automatic information retrieval on algorithms for detecting the topic of a text and on visualizing relationships;
- affective analysis, which provides insight into the personality of the sender and the nature of the relationship between sender and receiver, and includes the detection of anger and insults;
- social analysis, which attempts to infer information about the role of sender and receiver and the nature of their relationship by coding messages, for instance on signs for sarcasm, formality, or fear as was done by Duchenaud (2002).

In this study, we combine elements from both the affective and social analysis approaches and include both the qualitative information in the headers for structural

analysis and the content of the messages for relational analysis to develop a more detailed and rich understanding of the nature of specific ties, and the information that flows between these ties. Similar approaches have been used by Pantelli (2003) in studying power relationships that appear in e-mails and Duchenaud (2002) in a study of whether e-mails can reinforce organizational structures. The main difference between these studies and the one presented here is their primary focus on the role of e-mails, while for us the e-mails were used primarily as a data source. Investigating how e-mails are used to build networks was secondary to studying, through e-mails, how firms build networks. Another difference between our study and most other e-mail-based studies is that we do not focus on complete or closed intraorganizational networks (Borgatti and Molina 2003) or personal social networks (Pew Internet Project 2002), but on egocentric interorganizational networks of a startup firm.

In another paper, Wakkee et al. (2006) already described the technical aspects of collecting, preparing, and coding company e-mails. In this paper, we concentrate on the operationalization and analytical aspects of modeling network developments with e-mails. Our method combines descriptive statistical content analysis based on factual data in the message and headers (to, from, subject, and date-fields) with qualitative analysis focusing on the actual meaning of the rich data in the messages and their attachments. Most notably, the approach presented here allows for a combination of qualitative descriptions and graphical and (descriptive) statistical representation of the findings.

Using e-mails to model social network development is not exclusive to an international context (e.g., Haythornthwaite 2001). Yet, we think it is most appropriate in this context. After all, international ventures use e-mails more frequently than face-to-face meetings or telephone to overcome physical distance and different time zones (Constant et al. 1996; Wijayanayake and Higa 1999). Hence, e-mails account for a significant share of the available information present in these ventures. To illustrate our method we draw from examples of a case study of a Dutch International New Venture¹ (INV). INVs are involved in multiple value-creating activities around the world within six years after foundation (Oviatt and McDougall 1994; McDougall and Oviatt 2005). We use the examples to highlight critical issues rather than to present a full analysis of the findings.

Data sources in international social network research

A social network is a set of individuals or organizations that are connected by socially meaningful relationships (Burt 1992; De Koning 1999). Social network researchers describe the patterns underlying the relationships; trace the flow of information and resources through them; and specify their effects on the network members (Wellman 1997; Garton et al. 1997). Social networks can be studied as a whole, i.e., including all members and their mutual relationships, or from the perspective of a single actor, the ego. Adopting such an ego perspective (Granovetter 1973; Garton et al. 1997) here, the social network refers to the ego and the

¹ For a complete analysis of this case study, we refer to Wakkee (2004).

constellation of all the actors with which it interacts directly; these are called the alters. The relationship between the ego and a specific alter is called a tie. These ties connect the ego indirectly all those network members to whom their alters are connected.

To study ego networks, several data sources have been used and discussed in the literature. Most typically, researchers use standard surveys or name generators to ask respondents to describe each member of their network including their social characteristics (age, gender) and the characteristics of the relationship (personal or professional, customer-supplier; e.g., Wasserman and Faust 1994; Greve 1995). Other data sources include interviews, archival records (Welch 2000) and more recently online communities and, as is done here, e-mails (e.g., Kling 1996; Krebs 2003). A complete review of all these data sources is beyond the scope of this study. However, we do want to point to several benefits of company e-mails in comparison to other data sources and explain why e-mails offer excellent opportunities for modeling social network development of INVs.

First, e-mails contain data that allow for analyzing developments in both the ego-network structure and the content of the ties. Second, unlike interview and survey data (but like archival records and websites), e-mails are created independently of the investigation. Therefore, the risk of including information in a socially desirable way (in the light of the study) will be limited. In interviews or surveys, respondents are likely to forget (to mention) some current or historic contacts (Huber and Power 1985; Marin 2004). Also, the account of what happened may be colored or distorted by later experiences (Nancarrow and Brace 2000). E-mails provide either factual information or explicit accounts based on the perception of the composer *at the time* of writing. Consequently, we do not have to rely on personal memories when using e-mails, and we can conduct a truly longitudinal analysis.

Third, e-mail allows for instant feedback, transmission of multiple cues, use of natural language, and personal focus (Sroull and Kiesler 1986; Haythornthwaite 2001; Pantelli 2003; Flynn 2005).² So, the information is often very rich and revealing, especially when chains of e-mails and responses are taken together (Wakkee et al. 2006). Additionally, many e-mail messages have attachments (e.g., contracts, statistical data, pictures, and presentations, audio and video files) that contain extra information (Fielding 2002). Therefore, e-mail messages and attachments enable us to determine *and* make sense of the nature of many ties and of the information that flows through these ties. Through interviews or surveys it would be impossible to gain insight into so many ties in similar levels of detail because it would be too demanding for the informants. Much of the information contained in the e-mails and the attachments might also be found in archives (e.g., Armstrong and Jones 1987; Welch 2000). Yet, Wakkee et al. (2006) found that e-mails often include explicit explanations about the attached documents that you would not find in archival records.

Finally, previous research shows that people use e-mail to share information they would not exchange at all through other channels like fax, telephone, or face-to-face

² It should be noted though that the personal focus of e-mails can vary considerably. Depending on whether an e-mail is personal and being sent to only one or a small group of individuals or impersonal and sent to a large number of e-mail addresses (e.g., spam).

communication (Sproull and Kiesler 1986; Pantelli 2003).³ In addition, unlike archival records and information expressed in interviews and surveys, e-mails (when collected in real time) have not been subjected to any form of selection that leads to a loss of information that is potentially relevant to the investigation. Therefore, the messages or attachments will contain information that the researcher could never find in any other data source.

There are several downsides to using e-mails as a source of data as well. Besides the practical issues related to data preparation (Wakkee et al. 2006), a potential challenge is gaining access to this type of data in multiple companies to allow for cross-case comparison. In terms of actual analysis, researchers also have to consider the difficulty of making sense of short messages that are exchanged in response to previous communication through other channels. When using e-mails, researchers should be aware of the ethical implications of their study (Pettigrew 1997; Borgatti and Molina 2005; Breiger 2005). People often exchange personal and confidential information through e-mail. The alters will most likely not be aware of their messages being scrutinized in research, and they have not explicitly agreed to cooperate. Unfortunately, for the data to make sense, the researchers must know who the informants are to establish a link between the composer and the receiver (Borgatti and Molina 2003). Making e-mails anonymous is therefore no option. To protect the anonymity of the ego and its alters, aliases are used in all our reports. Also, we excluded the content of the messages containing a disclaimer from our analysis, using only the information in the headers in such cases.

Case study background and data

To explicate our method we draw on a case study of a Dutch INV with the alias Sound Inc. Sound Inc. was founded officially in 1998 by Mr. Wide and Mr. Path, but it had its origins in a scientific breakthrough made by Mr. Wide in 1994. Since then, he conducted further research and took preparatory steps toward commercializing his discovery through talking with potential users, financiers, and R&D partners around the world. Today, Sound Inc. develops and sells a range of high-tech sound-measuring sensors to universities, multinationals, and other organizations around the world. By December 2004, Sound Inc. employed five engineers. Interestingly, all of them worked only part-time: Mr. Wide had a “side job” as a university researcher, Mr. Path was the owner-manager of a second venture, and the employees were university students. So, although sales levels were growing, the venture remained small.

Clearly, this small size affected the size of the network and the number of interactions. When starting this investigation we did not foresee such limited growth, but the small size did make our case study feasible. After removing spam, personal messages and messages with disclaimers, our dataset included only 795 messages. Of these e-mails, 708 were collected in real time between December 2001 and May 2002,

³ This is probably the largest difference compared to using paper correspondence as a source of data (i.e., Thomas and Znaniecki 1918 and their classical case study on the Polish Peasant in Europe and America).

covering a period of roughly 6 months. The remaining 87 e-mails were archived messages sent and received between January 1995 and May 2001. These messages were collected post hoc from a backup file we obtained in June 2001; it should be noted that these messages had been subject to some selection process. Unfortunately, it is unclear what percentage of all messages sent and received in this period was lost in this selection process. E-mails exchanged with alters were coded as *external* ($n=603$) and e-mails exchanged between company officials were coded as *internal* ($n=192$). These groups were used differently in our analysis, as we describe in the following section. For practical reasons, we did not distinguish between different company officials when analyzing external e-mails, but treat them as one.

Because triangulation is an important part of any case study method (Eisenhardt 1989; Yin 1994), we used a large number of other sources such as interviews and archival records (see Table 1) to check our e-mail findings; however, a detailed description of these sources and how we used them is beyond the scope of this research.

Data analysis

We conducted two sets of analyses, the first using only the e-mails, the second using all the data sources as listed in Table 1. This allowed us to determine if the e-mails could be used as (1) a stand-alone data source, (2) if they only provided sufficient relevant information when combined with other sources in a triangulation process (Yin 1994), or (3) if the e-mails provided too little information to be of any value when analyzing specific network measures. Further, we incorporated quantitative and qualitative elements, taking mainly a deductive approach but combining it with a more inductive method (Denzin and Lincoln 2000). Although we used predefined categories derived from theory to guide our analysis (the deductive approach; Pettigrew 1997), we added codes and categories when the data suggested this was necessary to capture the “reality” of the company. Using such predefined lists of

Table 1 Overview of data sources

Data sources	Number
E-mail (real time)	708
E-mails (postdate)	87
attachments	78
Interviews	
company officials	6
external informants	3
Websites	
Sound Inc.	73
Other	38
Articles about Sound Inc. in Popular Press	6
Archived documents	
Faxes	234
Contracts	16
Memo's	21
Web statistics	5

categories, we analyzed the data through automatic and manual searches with the aid of a qualitative software package called NVivo (Richards and Richards 1995).

In particular, we combined the external e-mails exchanged with each particular alter into e-mail chains, starting with the earliest e-mail and ending with the most recent, and then incorporating passages about the particular alter from internal e-mails according to date. We then analyzed the conversations unfolding in these chains, and tapped into the richness of the data to make sense of how particular ties developed and were likely to develop in the future. Obviously, given the absence of a real conversation, this was only possible when chains consisted of multiple e-mails (e.g., “Dear Mr. Wide, Could you please send me a copy of your book? Regards Charles”; “Dear Charles, Of course, here it is. Best Wishes Alex Wide”). When multiple e-mails were available, reading the e-mail chain felt like being present at meetings or eavesdropping on phone calls. Due to the high frequency of some e-mail interaction, some conversations resembled intellectual midwifery (Groen and Nooteboom 1998).

Network variables

In the social network literature (e.g., Burt 1992; Wasserman and Faust 1994; Hansen 1995, Bruderl and Preisendorfer 1998; Rodan and Galunic 2004; Hite 2005) many different network variables have been studied. We will not include all these variables, but rather select some as illustrations of our arguments. In the following paragraphs we describe our approach to measuring our variables and findings in more detail and explain where difficulties and limitations were found and where e-mails offer extra inside in comparison to other data sources.

Network size, growth and frequency of interaction

The first step in our analysis was to conduct a basic structural analysis using only the *headers* of the external e-mails in our database. Following Krebs (2003) and Johnson (2003) we compiled a list of all the e-mail addresses listed in the “to”, “from”, “cc”, and “bcc” fields. Only e-mails addressed to (a small group of personally addressed) individuals were used, disregarding e-mails addressed to distribution lists. This first analysis allowed us to determine the *size* of (the e-mail) network and provided an indication of the *frequency* of the e-mail-based interaction for the entire network and for the individual alters. Further, by looking at the e-mails exchanged in specific periods of time, we could also determine the *growth* of the network.

As pointed out by Wijayanayake and Higa (1999), people make decisions regarding their choice of medium based on individual, contextual, and social influences. Even when people are geographically distant, they will never rely on e-mail alone for their communication. This means that if we want to develop a more complete picture of the development of the network we also need to look for information about the use of other media. Therefore, we conducted a secondary analysis based on the e-mail message. We manually inspected all external and internal e-mails for: (1) reference to (face-to-face or telephone) interaction with alters with whom no e-mails are exchanged and (2) reference to interactions through other

media (face-to-face, phone) with previously identified alters: "...to come back to our phone call from last Friday", or "...as discussed in our meeting from the 12th of March." The first allowed us to make a more accurate judgment of the size (and growth) of the network. The second enabled us to make more accurate judgments regarding the frequency of interaction.

Because some alters might use multiple e-mail addresses and multiple alters could work for the same organization, we proceeded with a detailed investigation of the e-mail addresses to establish any correspondence and to make corrections accordingly. Thus, we ended up with 168 individual alters, representing 150 organizations that were considered in our investigation of the other network variables.⁴ Thus, only a handful of interorganizational relationships are based on more than one personal relationship. We will come back to this issue in more detail in the discussion section.

To complete our analysis of network size, growth, and frequency of the interactions, we considered our other data sources. To begin with, we asked Mr. Path to think of alters with whom they never exchanged e-mails. Asking only one informant for this information was sufficient in this case. Because of the small size of the venture, the open flow of information between company staff (for instance, each sent or received e-mail is always sent as a carbon copy to all other company officials) and Mr. Path's central position in the organization as the founder and general manager, he was extremely well-informed regarding all external communication. He volunteered seven alters, four of which we had already identified. We also asked Mr. Path to indicate how often, on average, Sound Inc. interacted with each alter. This clearly raised the frequency of interactions from what we found from our e-mail investigation. Generally, we found a pattern: the more frequent the e-mail communication, the more frequent the communication through other channels. Finally, we combined this information with a qualitative content analysis of the archival records. This analysis yielded another small group of 16 alters in 12 organizations. Most of these ties were dormant and originated from the period 1994–1998 when the company was not even officially founded and when e-mail was not yet common.

Network heterogeneity

To gain insight into the degree of heterogeneity in the social characteristics of the alters, we measured global diversity (following Harveston 2000 and Hordes et al. 1995. this was measured using the number of countries where individual alters were principally located⁵); and the *range of organizations* they represented. To this end, we first inspected the e-mail addresses: these very often included their organization's name. If we recognized this (in the case of 53 of the 168 alters) we immediately listed the alter's location and their organization's nature (although still checking against other data sources as we continued).

⁴ The additional alters that we identified from the triangulation with the other data sources were not included in the analysis of the other variables as clearly the e-mails would not provide any information that could help in measuring these variables.

⁵ In our case many alters were multinational companies themselves. If this were the case, we listed the country where the individual alter was located rather than the location of the MNE's headquarters.

If we did not recognize the organization, we used valuable clues incorporated in the address: for instance, organizations with “.ac.uk”—addresses were classified as universities from the United Kingdom, while “.mil”—addresses were listed as governmental bodies (the army) of the U.S.A. This allowed us to classify the location and nature of organization of another 28 alters partially (location *or* nature) and of 12 alters completely (location *and* nature). When the information contained in the e-mail address was insufficient, a full automatic content analysis of the message was undertaken using a list of all countries (and their abbreviations) and of a wide range of synonyms for the organization types in Dutch, English, French, and German. Finally, we conducted a manual inspection of the e-mail messages to determine the accuracy and completeness of the automatic coding process. Checks with the key informants showed that this increased our success rate and enabled us to classify the nature of the organization of almost 85% of the alters and the country location of 92% of the alters. The quantitative findings are summarized in Table 2.

A qualitative investigation revealed that the majority of the domestic alters were established early on in Sound Inc.’s development (1994–1998). In this period, we also find a large number of international contacts, most notably from the USA, the UK, France, Sweden, and Japan, where some of the most dominant market players (possible competitors or alliance partners) are located. From 2000 onward, USA, Brazil, Australia, and Germany are important alter locations. The alters that could not be classified were typically alters with which only one or two e-mails were exchanged and with whom no real relationship developed. Regarding the range of organizations, our qualitative analysis showed that from the start both companies (mostly widely known multinationals) and universities were important alters. This is not surprising considering the high-tech and international nature of the industry in which Sound Inc. operates. The consultants with whom Sound Inc interacts come into view around the startup period (1998), whereas all the sales agents are found from 2000 onward.

Table 2 Quantitative results heterogeneity analysis

Variables	Measure	Results
Geographic diversity:	# countries in which the different alters are principally located (from country list)	Known: 155 ($N=168$) Domestic: 37 International: 118 from 37 different countries Unknown (International) 13
Range of organizations	# of different types of organizations with whom Sound Inc. interacts	Range = 7 Known: 143 ($N=168$) Company : 64 Research institute: 40 Government/public: 9 Conference organization: 8 Consultant: 11 Indep. Sales agent: 8 Other: 3 Unknown: 25

Nature of the relationship

To investigate the nature of individual ties, we sought to determine the *nature of the relationship*, the *strength of the tie*, and the *origin of the tie*. For all of these variables, we first conducted a more quantitative analysis of the e-mail chains using predefined codes and categories and counting the occurrences (see Table 3). As expected, we could only gain insight into the nature of those ties that were somewhat developed and for which we had multiple e-mails and/or other data sources. This was the case for only 93 alters. For the remaining 75 alters, the e-mails provided insufficient information. For instance, the information exchanged in the e-mail messages was limited to a request for *a copy of the Handbook of Sound Inc.'s core technology*. This book could be useful to students, researchers, potential R&D partners, customers, or others. Interestingly, we found that the relationship with several alters could be classified in more than one way: particularly, one university was classified as simultaneously the parent organization, supplier (of office space and R&D facilities), research partner, and network partner. Also, in 13 cases an R&D partner became a customer as well. We will come back to this issue later in the discussion.

The *strength of the ties* was measured as a combination of *frequency*, *multiplexity*, and *emotional bond* (Granovetter 1973). Above, we have already explained how we established the *frequency* of the interactions. To make more sense of these frequencies, we constructed timelines for each tie and listed each interaction as a dot on this timeline (see Fig. 1). This revealed that the interactions were not evenly distributed over time.

Closer inspection of the e-mail chains revealed that short periods of intensive, frequent communication (sometimes up to 10 messages were sent and received within 3 days) were followed by weeks or even months without any communication, with the peaks occurring around the closing of a sale (e.g., terms of sale, shipment details) and during the development of new types of sensors (when many technical details were discussed between the venture and, for instance, research institutes, or distributors).

To measure the *multiplexity*, we considered the *nature of the relationship* to see if one alter was listed under multiple categories. Further, we examined the number of *topics* (adopted from a list compiled by Van Der Veen 2004) discussed in the e-mails. If e-mails were solely a means of information exchange, then online ties

Table 3 Quantitative results for nature of the relationship

Variables	Results from email analysis only
Nature of the relationship:	Customer = 48 Supplier = 26 Parent organization = 1 Distributor = 31 R&D partner = 28 Network partner = 12 Other = 14 Unknown = 75 Total >168 due to double classifications

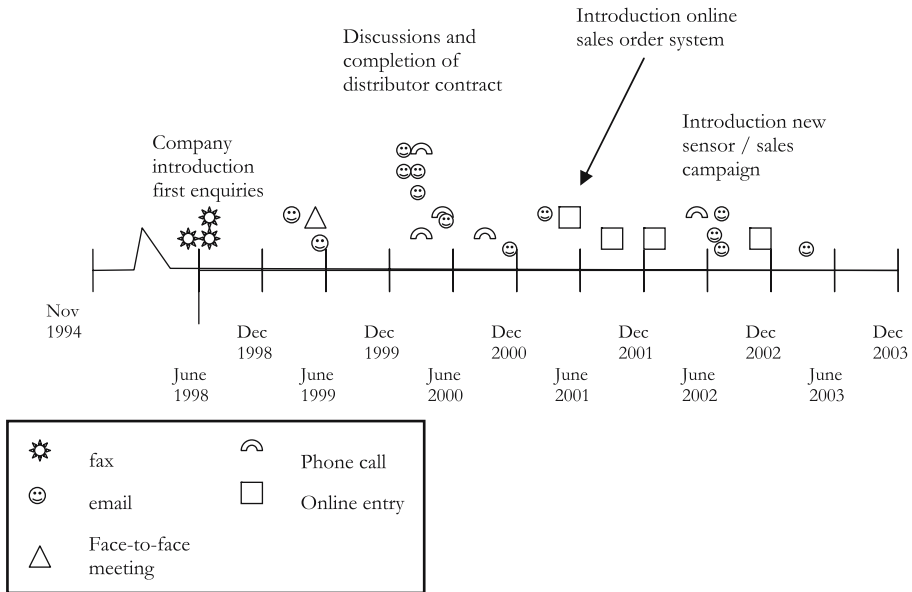


Fig. 1 Interaction timeline for distributor actor 1

would be narrowly specialized and aimed at finding social and other resources (Wellman 1997). Yet, as suggested by Hiltz et al. (1986), Walther (1994), through e-mail, ties exchange information; provides companionship, a sense of belonging, and (moral) support; and arranges services. Using our list of categories, we identified an average of 7.7 different topics discussed between the ego and its alters (typically related to two different value-added activities). As is shown in Table 4, the range (from a minimum of one topic to a maximum of 22 topics concerning four value-added activities) suggests that for Sound Inc., e-mail is indeed a way to develop both multiplex and focused, specialized relationships.

In comparison, only the archived faxes dating from 1994 to 1998 provided equally detailed information about the range of topics discussed. Yet, we found that on average, only 3.2 topics were discussed by fax. We have two explanations for this: First, in this period Sound Inc. was not yet officially founded and fewer activities had been initiated; second, owing to its informal nature, people will exchange information more easily by e-mail than through other media (Pantelli 2003). Probably, a combination of these two explanations describes the situation most accurately.

Table 4 Multiplexity findings

Codes	Finding
Average number of topics	7.7
Average number of value-added activities	2
Minimum number of topics	1 (n=31) (request for technical information)
Maximum number of topics	22 (n=2)
Maximum number of value-added activities	4 (n=2) (R&D, production, sales, services)

To determine the *emotional bond* between the ego and its alters, we adopted largely a qualitative approach. First, we started reading the salutation as we expected these might provide some interesting clues: e.g., compare the following quotes “Dear Dr. Hansen,..., With kind regards, A. Petersson” to “Pete, Cheers Alex.” Because we realized that the informal language often used in e-mails (e.g., Pantelli 2003) might be misleading and because different levels of formality are used in different cultural settings, we also searched for other clues in the conversations to infer the presence of close bonds. Sometimes these clues were very explicit: “Pete is a close friend of mine and I want to do him a favour.” Yet, such examples were relatively rare (only 12 examples). Further, from reading and rereading the e-mail chains, we developed an “intuitive” idea of whether bonds were emotionally close or more distant. Triangulation suggested that these intuitive ideas were fairly accurate, but more research is needed to draw explicit conclusions on this topic.

By combining the findings with respect to frequency of interaction, multiplexity of the relationship and the closeness of ties, we drew conclusions about the strength of the ties (Granovetter 1973; Bruderi and Preisendorfer 1998). Based on our observations, we decided to use a more detailed list of categories than the commonly used “weak” and “strong” ties and tried to classify the ties as: (1) new, (2) weak, (3) medium, (4) strong, or (5) inner circle. A check with the company informants revealed that the classification of 56 ties was similar to their own perception of the strength of their ties and somewhat similar in 24 cases.

Finally, we examined the origin of the ties. To establish *the origin* of the different ties, we measured *the date of the first interaction* and *the circumstances or location of the first interaction*. To determine the date of the first interaction, we listed the dates of the very first e-mail exchanged between the ego and the different alters. Next, we analyzed all the messages for three types of information: (1) explicit statements of introduction, indicating that an e-mail is in fact a first point of contact between the ties such as “Dear Mr. Wide, let me introduce myself, my name is ... Browsing the Internet, on your internet-site I found that you are developing sensors that just may do the trick”; (2) references to a first interaction in an e-mail between the ego and the particular alter: “Hi Alex, remember me, we met during the DAGA in July and discussed the possibilities of a dealership...”; or (3) references to the origin of tie in an e-mail to a colleague or third party: “Tom and I went to college together.”

This way, we could determine the origin (date, circumstance, or both) of less than 50% of all ties (see Table 5). Nevertheless, we consider this sufficient to consider e-mail a relatively good source for determining tie origin when compared to other sources. Many alters for which we could not determine the origin had been established in the period (1995–1998) before e-mail was used commonly. For ties originating after 1998, the success rate was considerably higher (almost 70%).

Discussion and conclusions

We have presented an innovative method for modeling the network development of INVs using e-mails as the main source of data. Our findings show that combining quantitative and qualitative approaches provides an opportunity to gain insight into

Table 5 Origin findings

Codes	Finding
Precise origin (timing and situation)	$n=27$ e.g., Dear Alex, We do not yet know each other, let me first introduce myself...our mutual supplier Microtech suggested you might be able to...
Reference to fist meeting on tradefair or conference without listing the precise date	$n=11$
Direct mentioning that firm was found online, yet it is unclear if a tip preceded the online 'search and find'	$n=16$
Introduced by third party, but the precise date of introduction is unknown	$N=7$

the development of the structure and composition of the networks of an INV, the nature of the different ties, and the content of interactions as these ties develop.

The main theoretical contribution to the field of international entrepreneurship and the understanding of new venture internationalization consists of gaining a better understanding, by studying the development of an INV's social network, of why and how such ventures are able to become globally active in a relatively short period of time. Previous studies on the role of networks have pointed to client followership (e.g., Bell 1995) or mentioned the role of domestic clusers (Zucchella 2001) as drivers of internationalization; others have pointed to the importance of visiting trade fairs to establish foreign ties (e.g., Rasmussen et al. 2001). In this study, we found detailed evidence of a variety of paths that lead to the formation of individual new international contacts. Most notable in this case were the use of brokerage, visiting trade fairs, and proactive use of the internet, both to identify prospective contacts and to be visible online to others. As such, the findings presented in the illustrative case study can be used to advice other startups in relation to international network formation.

With regard to structural variables, we found it relatively easy to measure a large number of variables (size, growth, frequency of interaction) through analyzing e-mails individually or combined in e-mail chains. When focusing on content-related variables, individual e-mails provided relatively little information; we could only make sense of the interaction between the ego and its alters when we began to analyze the conversations that unfolded in e-mail chains.

Besides the structure and content of the relationship, network development over time also became apparent from analyzing the e-mails as chains of conversations *over time* as called for by Pettigrew (1997) and Coviello and Jones (2004). In particular, we examined if particular ties became stronger (e.g., more or less frequent interaction, increasing number of conversation topics and developing emotional bonds) by analyzing the content of the e-mail chains. We examined if the network grew in size and in heterogeneity by breaking up the complete timeframe covered by this study into 3-month periods and comparing the findings for these periods, and by looking for trends developing from one period to the next.

The quality of a method is not only dependent on the range of variables that can be measured with it, but also on the accuracy of the findings. Our research shows triangulation certainly improves the accuracy of the findings. When using only e-mails

the quantitative findings should therefore be interpreted as “minimal” or “at least” findings because most interactions through other media would be missed. For instance, we might observe that the ego interacts with a particular alter on a biweekly basis through e-mail, whereas in fact the interaction takes place twice a week by phone and e-mail combined. Despite this, we conclude that compared to other data sources e-mails provide more qualitative insights and richer details regarding the network interaction than other sources, allowing for better sensemaking of descriptive variables.

In reflecting upon our study, it is important to realize that the illustrative case study was set in a high-tech emerging industry. As a consequence, the number of players (companies, individuals, universities, etc.) that were active in the field at the time of investigation was rather limited. This clearly affected the outcomes of the study. The number of ties was limited, making it easier to conduct an in-depth investigation. Also, the developing nature of the industry affected the nature of the ties and the information that was exchanged through e-mail. As mentioned previously, many ties were classified as multiplex relationships: for instance, the university acted as a knowledge and network broker, provided access to research facilities, and could be considered a source of legitimacy, while many companies were both clients and co-developers. In a more mature industry, the specialization of tasks and the level of competition are likely to have reduced the amount, nature and quality of the information being exchanged through e-mail, as parties may become more secretive.

The usefulness of a method is also dependent on practical considerations. In general, qualitative analysis is very time-consuming (Yin 1994) and e-mail analysis even more so because of the need to create and disentangle e-mail chains, and to resolve language and typographical difficulties (Wakkee et al. 2006). We feel, however, that researchers should be able to overcome such problems with good planning and preparation and possibly by cooperating with others. One of the main benefits of using e-mails is that the researcher does not have to be present at the site, nor does it require considerable time from the informants. When compared to other data sources, it seems that only direct questions in interviews regarding the origin of previously identified ties might provide more information, but establishing the origin of all 168 contacts in this way would be far too time consuming for the informant, thus making this a poor alternative. This means that for longitudinal, in-depth analysis of social network development using e-mails, though not accurate and complete, offers possibly the best opportunities for researchers.

Challenges and needs for further research

The methods and illustrations of the findings presented in this paper are based on a single case study to model social network developments in INVs. As discussed above some of the outcomes of the illustrative case study are likely to be attributable to the specific context in which the venture developed: a high-tech emerging industry, surrounded by knowledge institutes and starting around the time that the use of e-mail was rising rapidly. Clearly, further case studies are necessary to build experience and fine-tune the method and to determine which other network dimensions can best be

measured using e-mails. One of the aspects that should be taken into consideration is the applicability of our method in different contexts. We expect that the international nature of Sound Inc.'s network positively affected the accuracy of our findings. Because of large distances and differences in time zones, e-mail was used more commonly than would have been the case for a domestic network.

Furthermore, analyzing company e-mails only allows for ego-level analysis of interorganizational networks. The boundaries of such networks cannot be determined and consequently it is impossible to gain access to all the e-mails of all actors. In closed networks conducting a full-scale structural analysis would be possible (Garton and Wellman 1995; Haythornthwaite and Wellman 1998), providing all network members are willing to grant the researcher access to their e-mail (archives).

Next, the method described here is particularly applicable to modeling network development of INVs because these firms commonly use e-mail to communicate with their geographically dispersed alters. Because they are young, network developments can be mapped from the earliest stages, and because they are typically small they produce a relatively manageable number of e-mails. When conducting research at larger and/or older ventures, researchers might consider analyzing the headings of all e-mails to obtain a general picture and then focus on a specific time frame or department for the in-depth analysis to keep the data set within reasonable size and thereby to keep the study feasible. Although a full discussion of the different sampling approaches is beyond the scope of this paper, it seems that combining statistical sampling with interviews seems to be a good starting point (Rothenberg 1995). Further research is, however, necessary to gain further insight into the minimal number of e-mails required to generate a fair grasp of the network (developments) in a company on the one side and on a maximum number of e-mails to keep the research within the limits of practical feasibility on the other.

During our investigation, we observed many interesting developments within the company in relation to company's culture, problems associated with new product development, formulation of new strategic goals and many more. This suggests that social networking is only one of the many topics related to entrepreneurship and new business creation that can be investigated with the use of e-mails as a source of qualitative data. Further research is necessary to explore these possibilities.

Our data suggested that in this case, the vast majority of the interorganizational relationships were based on relationships between the ego and a single individual. This raises questions concerning the institutionalization of these relationships. Unfortunately, the limited time span of our investigation does not allow us to determine if the relationships are based on personal or organizational connections (e.g., Porter 2006). An interesting opportunity for future research would therefore be to examine what happens if the individual ties leave their current organization: does the relationship between the ego and the alter organization remain intact, or does it transfer with the individual tie to its new organization, or both? Studying the effects of such transitions can help us develop a better understanding of how networks of startups come into existence and how they develop over time. Studying such effects would, however, involve studying e-mails over a longer period of time. Studying only the "to" and "from" fields may be required to identify potential shifts as argued by Donath (2004), but to make sense of the underlying mechanisms would again require content analysis of the e-mail bodies.

Of course, the limitations of this data source need to be kept in mind. Regarding the specific network variables, it is obvious that some are easier to measure using e-mail than others. In particular, further research would be necessary to gain a better understanding of the extent to which we can measure emotional bonds from e-mails, as the approach we adopted was rather subjective and intuitive in nature. Another constraint concerns the requirements of ethics committees. In the Netherlands, no formal ethical guidelines exist that forbid the use of e-mails without asking explicit permission from all the network actors. We realize that in some countries asking explicit permission is mandatory. This might lead to unworkable situations. In these cases, using e-mails only to identify network structures following the method proposed by Krebs (2003) and Johnson (2003) might offer a less detailed yet still valuable alternative.

To conclude, in our study and others it is clear that INVs use e-mail for the majority of their interactions with globally dispersed alters and also for a considerable share of their internal communication. A lot of the information about network development and company development in general is thus stored in company e-mails. Therefore, if we want to conduct in-depth case studies, e-mails should simply not be overlooked.

Acknowledgments We are grateful for the extensive and valuable comments on earlier versions of this paper of the anonymous reviewers of the *Journal of International Entrepreneurship* and like to thank Aard Groen, Peter Groenewegen, and Paul Kirwan for their suggestions. Their comments have helped us to improve the arguments in this paper as well as to solve language and style issues.

References

- Andersson S, Victor I (2003) Innovative internationalisation in new firms: born globals- the Swedish case. *Int J Entrep* 1(3):249–276
- Armstrong J, Jones S (1987) *Business documents: their origins, sources, and uses in historical research*. Mansell, London
- Bell, J (1995) The internationalization of small computer software firms—a further challenge to “Stage: Theories.” *Eur J Mark* 29(8):60–75
- Bell J, McNaughton R, Young S (2001) “Born-again Global” firms: an extension to the ‘Born Global’ phenomenon. *J Internat Manag* 7:173–189
- Bell J, Loane S (2002) Entrepreneurship research in Europe: innovative methods in the exploration of internationalisation issues. Paper presented at the First European Summer University: Entrepreneurship Research in Europe: Specificities and Perspectives, Sept. 19–22, Valence (France), INPG/ESISAR
- Blomstermo A, Sharma D (2002) Born global: a study of learning in networks. Paper presented at the 28th EIBA conference, Athens, Greece
- Borgatti SP, Molina J-L (2003) Ethical and strategic issues in organizational network analysis. *J Appl Behav Sci* 39(3):337–350
- Borgatti SP, Molina J-L (2005) Toward ethical guidelines for network research in organizations. *Soc Netw* 27(2):107–117
- Breiger RL (2005) Introduction to special issue: ethical dilemmas in social network research. *Soc Netw* 27(2):89–93
- Bruderl J, Preisendorfer P (1998) Network support and the success of newly founded businesses. *Small Bus Econ* 10:213–225
- Burt RS (1992) *Structural holes, the social structure of competition*. Harvard University Press, Cambridge
- Constant D, Sproull L, Kiesler S (1996) The kindness of strangers: the usefulness of electronic weak ties for technical advice. *Organ Sci* 7(2):119–135
- Coviello N, Munro H (1995) Growing the entrepreneurial firm: networking for international market development. *Eur J Mark* 29(7):49–61

- Coviello N, Jones M (2004) Methodological issues in international entrepreneurship research. *J Bus Venturing* 19(4):485–508
- De Koning A (1999) Conceptualizing opportunity formation as a socio-cognitive process. Dissertation, INSEAD Fontainebleau
- Denzin NK, Lincoln YS (2000) The discipline and practice of qualitative research. In: Denzin NK, Lincoln YS (eds) *Handbook of qualitative research*, 2nd edn. Sage, Thousand Oaks, CA, pp 1–28
- Dimitratos P, Jones MV (2005) Future directions for international entrepreneurship research. *Int Bus Rev* 14(2):119–128
- Donath J (2004) Visualizing email archives. Sociable Media Group Working Paper, accessed on 17 July 2006 at <http://smg.media.mit.edu/papers/Donath/EmailArchives.draft.pdf>
- Ducheneaut N (2002) The social impacts of electronic mail in organizations: a case study of electronic power games using communication genres. *Inf Commun Soc* 5(2):153–188
- Easton G (1995) Methodology and industrial networks. In: Möller K, Wilson DT (eds) *Business marketing: an interaction and network perspective*. Kluwer Academic Publishing, Norwell, MA, pp 411–491
- Eisenhardt KM (1989) Building theories from case study research. *Acad Manage Rev* 14(4):532–550
- Fielding N (2002) Automating the ineffable: qualitative software and the meaning of qualitative research. In: May T (ed) *Qualitative research in action*. Sage, London, pp 161–178
- Flynn DA (2005) Seeking peer assistance: use of email to consult weak and latent ties. *Libr Inf Sci Res* 27(1):73–96
- Garton L, Wellman B (1995) Social impacts of electronic mail in organizations: a review of the research literature. *Communication Yearbook* 18:434–453
- Garton L, Haythornthwaite C, Wellman B (1997) Studying online social networks. *Journal of Computer-Mediated Communication* 3(1), retrieved April 12th 2005 from <http://www.ascusc.org/jcmc/vol3/issue1/garton.html>
- Granovetter M (1973) The strength of weak ties. *Am J Sociol* 78(6):1360–1380
- Greve A (1995) Networks and entrepreneurship: an analysis of social relations, occupational background, and use of contacts during the establishment process. *Scand J Manag* 11(1):1–24
- Groen A, Nooteboom B (1998) Environmental innovation: knowledge and networks. Unpublished paper
- Halinen A, Törnroos JA (2005) Using case methods in the study of contemporary business networks. *J Bus Res* 58(9):1285–1297
- Hansen EL (1995) Entrepreneurial networks and new organization growth. *Entrep Theory Pract* 19:7–20
- Harveston PD (2000) Synoptic versus incremental internationalization: an examination of born global and gradual globalizing firms. Unpublished Doctoral Dissertation, University of Memphis
- Haythornthwaite C, Wellman B (1998) Work friendship and media use for information exchange in a networked organization. *J Am Soc Inf Sci* 46(12):1101–1114
- Haythornthwaite C (2001) Tie strength and the impact of new media. Proceedings of the 34th Hawaii International Conference on System Sciences. IEEE Computer Society
- Hiltz SR, Johnson K, Turoff M (1986) Experiments in group decision making: communication process and outcome in face-to-face versus computerized conferences. *Human Commun Res* 13:225–252
- Hite JM (2005) Evolutionary processes and paths of relationally embedded network ties in emerging entrepreneurial firms. *Entrep Theory Pract* 29(1):113–144
- Hordes MW, Clancy JA, Baddaley J (1995) A primer for global start-ups. *Acad Manage Exec* 9(2):7–11
- Huber GP, Power DJ (1985) Retrospective reports of strategic level managers: guidelines for increasing their accuracy. *Strateg Manage J* 6:171–180
- Johnson S (2003) Emerging technology, pass your email through some new software and the answer will become obvious. *Discover Magazine* 24(04) <http://www.discover.com/issues/april-03/departments/feattech>
- Jones MV, Dimitratos P (2004) Emerging paradigms in international entrepreneurship: a synopsis. In Jones MV, Dimitratos P (eds) *Emerging paradigms in international entrepreneurship*. Edward Elgar, Cheltenham, pp 3–18
- Kling R (1996) Social relationships in electronic forums: hangouts, salons, workplaces and communities. In: Kling R (ed) *Computerization and controversy: value conflicts and social choices*, 2nd edn. Academic Press, San Diego, CA pp 426–454
- Krebs V (2003) Data Mining Email to Discover Social Networks and Emergent Communities, <http://www.orgnet.com/email.htm>
- Marin A (2004) Are respondents more likely to list alters with certain characteristics? Implications for name generator data. *Soc Netw* 26:289–307
- McDougall PP, Shane S, Oviatt BM (1994) Explaining the formation of international new ventures: the limits of theories from international business research. *J Bus Venturing* 9(6):469–487

- McDougall PP, Oviatt BM (2005) Fundamentals of international entrepreneurship. *Entrep Theory Pract* 29(5):537–554
- Nancarrow C, Brace I (2000) Saying the right thing: dealing with socially desirable responses. *Bristol Business School Teaching and Research Review* 1(3):1468–1478
- Neergaard H, Ulhøi J (eds) 2006 *Handbook of Qualitative Research Methods in Entrepreneurship Research*. Edward Elgar, Cheltenham
- Oviatt BM, McDougall P (1994) Toward a theory of international new ventures. *J Int Bus Stud* 25(1):45–64
- Oviatt BM, McDougall PP (1995) Global start-ups: entrepreneurs on a worldwide stage. *Acad Manage Exec* 9(2):30–44
- Pantelli N (2003) Richness, power cues and email text. *Inf Manage* 40(2):75–86
- Pettigrew AM (1997) What is a processual analysis? *Scand J Manag* 13(4):337–348
- Pew Internet Project (2002) *The Internet Goes to College*, Report no 71, 15 September
- Podolny JM, Baron JN (1997) Resources and relationships: social networks and mobility in the workplace. *Am Sociol Rev* 62:673–693
- Porter KA (2006) A myriad of networks: understanding the relationships between organizational and entrepreneurial team networks. Paper presented at the Strategic Entrepreneurship Workshop, Vrije Universiteit, 3–4 July
- Rasmussen ES, Madsen TK, Evangelista F (2001) The founding of the born global company in Denmark and Australia: sensemaking and networking. *Asia Pac J Mark Log* 13(3):75–107
- Rialp A, Rialp J, Knight GA (2005) The phenomenon of early internationalizing firms: what do we know after a decade (1993–2003) of scientific inquiry? *Int Bus Rev* 14(2):147–166
- Richards L, Richards T (1995) Using computers in qualitative research. In: Denzin N, Lincoln Y (eds) *Handbook of qualitative research*. Sage, Thousand Oaks, CA, pp 445–462
- Rodan S, Galunic C (2004) More than network structure: how knowledge heterogeneity influences managerial performance and innovativeness. *Strateg Manage J* 11(1):541–562
- Rothenberg RB (1995) Commentary: sampling in social networks. *Connections* 18(1):104–110
- Sharma DD, Blomstermo A (2003) The internationalization process of born globals: a network view. *Int Bus Rev* 12:739–753
- Sproull L, Kiesler S (1986) Reducing social context cues: electronic mail in organizational communication. *Manage Sci* 32:1492–1512
- Thomas WI, Znaniecki FW (1918) *The Polish peasant in Europe and America: monograph of an immigrant group*. New York
- Van der Veen M (2004) *Explaining e-business adoption: innovation and entrepreneurship in Dutch SMEs*. University of Twente, Enschede
- Wakkee IAM (2004) *Starting global: an entrepreneurship-in-networks approach*. Doctoral Dissertation, University of Twente, The Netherlands
- Wakkee IAM, Danskin P, During WE (2006) Using email as a source of qualitative data. In: Neergaard H, Ulhøi J (eds) *Handbook of qualitative research methods in entrepreneurship research*. Edward Elgar, Cheltenham, pp 331–359
- Walther JB (1994) Anticipated ongoing interaction versus channel effects on relational communication in computer-mediated interaction. *Human Commun Res* 20:473–501
- Wasserman S, Faust K (1994) *Social network analysis: methods and applications*. Cambridge University Press, New York
- Welch C (2000) The archaeology of business networks: the use of archival records in case study research. *J Strat Mark* 8:197–208
- Wellman B (1997) An electronic group is virtually a social network. In: Kiesler S (ed) *The culture of the internet*. Lawrence Erlbaum, Hillsdale, NJ, pp 179–205
- Wijayanayake J, Higa K (1999) Communication media choice by workers in distributed environments. *Inf Manage* 36:329–338
- Yin R (1994) *Case study research: design and methods*, 2nd edn. Sage, Thousand Oaks
- Zucchella A (2001) Born global versus gradually internationalizing firms: an analysis based on the Italian case. Paper presented at the 28th EIBA conference, Athens, Greece