Identifying Information Security Risks in a Social Network Using Self-Organising Maps

Rudi Serfontein(✉), Hennie Kruger(✉), and Lynette Drevin(✉)

North-West University, Potchefstroom, South Africa
{rudi.serfontein,hennie.kruger,
lynette.drevin}@nwu.ac.za

Abstract. Managing information security risks in an organisation is one of the most important tasks an organisation has. Unfortunately, due to the complexity of most organisational systems, identifying information security risks can be difficult. One way to identify possible risks in an organisation is to make use of Social Network Analysis (SNA). While they can be used to identify risks, the metrics calculated using SNA are often numerous and daunting to managers unfamiliar with SNA. Furthermore, as the data in this form tend to be uncomfortable to process, educating managers about risks in their organisation can be quite difficult. Also, as these metrics often require quantitative processing in order to be useful, SNA on its own is not always an attractive method to use to identify risks in an organisation. In this paper the use of self-organising maps to identify possible information security risks in an organisation is investigated. Risk data were obtained from an organisation that deals in risk management, which were used to build a social network. A number of metrics associated with risk were calculated from the network, and these metrics were used to cluster the various entities using a self-organising map. Certain entities that pose a possible information security risk were identified. The results suggest that it may be viable to use self-organising maps, in concord with SNA, to more easily identify risks in an organisation using visual methods.

Keywords: Self-organising maps · Social network analysis · Information security

1 Introduction

Information security risk management is one of the most crucial parts of information security and should be one of the most important actions taken by organisations [1]. Unfortunately, due to the relative complexity of most organisational systems, identifying information security risks that are inherent to people using the systems, and making managers aware of them, is often quite difficult. One of the methods proposed

Published by Springer Nature Switzerland AG 2019
L. Drevin and M. Theocharidou (Eds.): WISE 2019, IFIP AICT 557, pp. 114–126, 2019.
https://doi.org/10.1007/978-3-030-23451-5_9

in recent years to address such risks involve the use of Social Network Analysis (SNA) [2–4]. SNA is a method that can be used to evaluate an organisation, for instance a community or business, in such a way that social interactions can be studied quantitatively, rather than qualitatively [5]. It does however have a significant drawback in that large networks, when visualised, may have so many nodes and arcs that the network is visually incomprehensible. In order to address this drawback, a number of studies have employed techniques that alter significant nodes and edges of a visualised network in order to draw attention to certain aspects. Some of these techniques include differentiating the colour of nodes and edges [6], using differing sized nodes to correspond to certain metrics [7], and using labels of various sizes [3]. A somewhat more novel technique makes use of Self Organising Maps (SOMs) to directly visualise network data [8]. A SOM is an effective technique that can be used not only to visualise high-dimensional data, but to visualise it in such a way that the result can act as both a similarity graph and a clustering diagram [9]. The SOM technique can be used to identify similar nodes within a social network, even in the presence of seemingly contradicting attributes, and present this data in a way that managers are quickly informed of risks in the network. SOMs have also been used in information security research to propose improvements to intrusion detection methods [10], and as a method for analysing information security behavioural data [11, 12]. While the approach suggested by Boulet, Jouve, Rossi and Villa [8] does allow for social networks to be visualised as SOMs, it has a shortcoming in that the SOMs generated can not necessarily be used in a way that is relevant to the process of identifying possible information security risks. This is mainly as a result of the fact that, in order to identify risks within a social network, a number of metrics calculated from the network data is used rather than the raw data itself.

In this paper the feasibility of using existing SOM techniques to inform managers of possible risks in an organisation is discussed. The value of such an application is twofold; firstly, by using a visualisation method that reduces the amount of data that is visualised, the often confusing graphs produced by traditional SNA visualisation techniques can be replaced with SOMs that are easier to process graphically. Furthermore, as SOM algorithms produce maps that naturally display data of interest, analysis and evaluation of the data should no longer require the visualisation results to be adapted (node enlargement, colouration, etc.) in order to be meaningful. Secondly, as SOMs organise similar data into clusters, their application should make it easier to inform a manager of groups of similar at-risk entities. This is thanks to the clustering done by the SOM algorithms, as, due to the relationship between certain SNA metrics and the CIA triad (Confidentiality, Integrity and Availability), as will be discussed in Sect. 2, clustering the nodes according to these metrics allows an evaluator to quickly identify similar problematic nodes. Furthermore, because the SOM algorithm uses the calculated SNA metrics as attributes to determine the clusters, the clusters themselves can be used to infer similarities that may not be readily transparent from the available data. Another advantage is that managers can be informed and educated of possible risks in an organisation early on, which may aid in developing effective awareness, education, and training programs. The graphical nature of SOM may also make it a useful tool for training inexperienced risk managers, and can potentially aid in identifying standard trends and patterns.

In the remainder of the article the background, research methodology, and results, will be discussed respectively. The background discussion will focus on SNA in the context of information security, as well as SOMs. The discussion of the method will focus on both the application of the techniques, and the data collection phase. The paper will then conclude with a discussion of the results and implications.

2 Background

The primary theme of this paper is the use of Social Network Analysis (SNA) metrics as inputs for a Self-Organising Map (SOM), which should aid in evaluating risk in an organisation. In order to demonstrate how this can be done, five SNA metrics will be discussed briefly. While there are dozens of SNA metrics that can potentially be used, the five discussed here were chosen based on their established relationships with risk in the literature. The section will start with a description of SOMs, followed by the evaluation of the selected SNA metrics in relation to risk.

2.1 Self-Organising Maps (SOM)

The self-organising map (SOM) is a neural network technique that can be used to visualise and evaluate high-dimensional data [13]. The SOM technique uses given data to produce a self-organising neural network wherein the data points are clustered into topographical regions [14]. This visualisation technique has a wide range of known applications, from evaluating comparable biological adaptations [15] and improving optimisation algorithms [16, 17], to clustering data for problem-solving purposes [14, 18]. One of the greatest advantages SOM has over other high-dimension visualisation techniques is that it produces a two-dimensional topographical map that can be evaluated and interpreted without any special knowledge or skills. In addition to clustering known data points, depending on the data, the technique can also be used in vector quantisation, and as a regression modelling technique [13]. All these methods can arguably be used to obtain valuable information about data, but in the context of this paper only the clustering function of SOM will be considered. The algorithm for developing a SOM [19] is shown below.

Input: Dataset *N*
Output: A topographical map *M* containing the data from *N*, sorted into topographical areas

Variables:
w_{ij} – Weight vector describing topographical area ij; either randomised or defined at start
x – An input vector contained in *N*
α – Learning rate that is a slowly decreasing function of time

```
while Stop condition is false
    For each $x$ in $N$
        For each vector $j$
            | Compute $D(j) = \sum_i (w_{ij} - x_i)^2$
        end
        Find index $J$ such that $D(J)$ is a minimum
        For all units $j$ in a topographical area $J$, and for all $i$:
            | Compute $w_{ij}(new) = w_{ij}(old) + \alpha[x_i - w_{ij}(old)]$
        end
    end
    Update $\alpha$
    Reduce radius of topographical area at specified times
    Test Stop condition
end
return $M$
```

This algorithm produces one map with all of the entities sorted into clusters. Certain software suites, such as Viscovery SOMine [20], provide additional information by colouring the same map using values from different attributes.

As stated, SOMs can be used to cluster high-dimensional data on a two-dimensional map, producing a result that can be interpreted easily without training. This makes the technique especially valuable to those in managerial positions, as these individuals may not have the time to study large reports and data sets in detail, and may also hold true for the outcomes of SNA based studies – especially if the resulting network is particularly large or complex. By calculating the SNA metrics, as discussed in the next section, and applying SOM to the resulting data set, the risks posed by certain individuals or groups can be determined and presented visually in a way that is easy to process and interpret. Additionally, a number of at-risk individuals may be identified that would not necessarily have been evident through the use of more traditional visualisation techniques such as bar-graphs. It is possible for an individual to have all the traits of a high-risk individual and not be an obvious risk from the data itself. In these instances, a clustering technique such as SOM can be used to identify individuals that have similar, possibly hidden, attributes. This makes it significantly easier to address certain information security risks, as larger scale programmes can be developed to target groups consisting of similar individuals.

In summary, SOM is a valuable technique to use in addition to SNA, as the clustering function of SOM can be used to infer invaluable information about information security risks if the correct and relevant SNA metrics are used.

2.2 Social Network Analysis (SNA) in the Context of Information Security

One of the most well-known frameworks for information security is the CIA triad [21], which references Confidentiality, Integrity, and Availability. ***Confidentiality*** describes the access rights that users have to a piece of information. For example, a manager having confidential access to certain business data that his employees do not, or should not, have. One possible SNA metric that can be used to evaluate a risk to confidentiality is *closeness centrality*. Closeness centrality is calculated by determining all the shortest distances to all other nodes within the network [22], so a node with a high closeness centrality has a large number of close relationships to other nodes in the network. Such a node may therefore have access to information that it should not have access to. Alternatively, if the node is an object or a resource, such as a shared computer or a photocopier, it could become a significant confidentiality risk if malware or untrustworthy maintenance personnel are involved.

Integrity describes not only how accurate any piece of information is but, by extension, how trustworthy it is. One of the SNA metrics that can be used to evaluate the risk a node poses to the integrity of the information in the network is *total degree centrality*. The total degree centrality measure is concerned with an individual node's position within the network [22, 23], and is determined by using the number of nodes leading into and out of a node. A node with a high total degree centrality is well connected within the network, and may have enough influence over other nodes to impact the integrity of the information passing through them. Consider, as an example, an office worker with a high total degree centrality that has to capture data for a corporate database. If this worker were to make a mistake in capturing the data, the integrity of the data that a large number of nodes rely on may be compromised.

The *betweenness centrality* measure can be used to identify nodes that are risks to both integrity and confidentiality. This measure is a representation of the number of times that a particular node is part of the shortest path between any two nodes in the network [24]. It is reflective of the number of indirect nodes that are connected to that node. To demonstrate the rationale behind using betweenness centrality as an indicator of risk to both integrity and confidentiality, consider a department with a "go-to" individual. This individual will likely have access to greater amounts of information than is ideal, and would be in a position to alter the information flowing through the network. Furthermore, as nodes with high betweenness measures tend to act as brokers, this individual may be seen as a trustworthy shortcut to obtain information in the network, which places it in a position to obtain greater amounts of information, as well as manipulate information as it flows through the network.

The final member of the triad, ***availability***, deals with the ability to access the data in a timely manner. Availability is often at odds with both confidentiality and integrity, as systems meant to protect availability and confidentiality often impact on the availability of the data. With regards to SNA, one of the metrics that may identify a high risk node in terms of availability is the one that identifies a node as a *boundary spanner*. A boundary spanner is a node that, if removed, will cause one part of the network to become completely isolated [25], thereby negatively impacting the availability of information in certain parts of the network.

The final SNA metric to be mentioned here is *eigenvector centrality*. Eigenvector centrality measures the extent to which a particular node is connected to highly connected nodes [22]. Nodes that have a high eigenvector measure are considered to possess emergent leadership properties [26] and may be considered a potential risk to confidentiality, integrity, and availability. For example, consider the impact an informal leader can have on information in a network. Confidential information may be shared with such a node as a result of the connections with highly connected nodes, whereas the integrity of the data in the network may be impacted by the additional knowledge the node obtains. Availability may also be impacted negatively if the emergent leader convinces other nodes to delay the flow of information, or if the information is redirected through the network along suboptimal routes. A summary of the SNA metrics discussed, and how they relate to the CIA triad, is given in Table 1.

Table 1. SNA metrics in the context of information security

SNA metric	CIA Triad			Rationale
	C	I	A	
Total degree centrality		X		Nodes with a high total degree centrality have influence in the network and are connected to a significant portion of the network
Closeness centrality	X			Nodes with a high closeness centrality have access to a significant amount of information in the network
Betweenness centrality	X	X		Nodes with a high betweenness centrality may be prone to information brokering and tampering
Boundary spanner			X	If a boundary spanner node is removed, an entire section of the network becomes isolated
Eigenvector centrality	X	X	X	Nodes with a high eigenvector centrality are considered emergent leaders and, depending on their influence and attitude, may be a general risk, depending on the circumstances

It should be emphasised that, while this brief discussion focussed on each member of the CIA triad individually, confidentiality, integrity, and availability are all interconnected. It is possible, for instance, for a significant enough increase in confidentiality to result in a significant reduction in availability. This is also true for the relationship between confidentiality and integrity, and the relationship between integrity and availability. The goal is typically to find a balance between these three aspects that is appropriate to the particular situation. This interrelatedness should be kept in mind when evaluating risks using SNA, as well as when selecting controls to address these risks.

From this short discussion it is clear that SNA metrics can be used to identify and possibly evaluate risks in terms of the CIA triad. The use of SNA metrics as input data for a SOM is therefore appropriate, and its use in this manner may help to identify and visualise risks in an organisation. This should help to improve overall awareness of risks in the organisation, aid in training managers, and may even help to determine overall preventative measures. The application in the rest of the paper uses the five SNA metrics discussed in this section as input for a SOM.

3 Method

The study was conducted using data provided by a manager from a large company that deals with risk evaluation. The data are confidential, and were subsequently anonymised prior to publication. Using this data, a network was built that describes the relationship between various entities. This network is shown in Fig. 1.

The entities, or nodes, of the network include 26 real-world risks, 612 controls, 6 risk owners, 26 control owners, 13 risk coordinators and 12 governing bodies. The risks are those risks that the organisation has to manage, whilst the controls are those controls used to manage the risks. The risk- and control owners are ultimately responsible for the risks and controls respectively, whereas the risk coordinators ensure that the correct risks are managed using the appropriate controls. The governing bodies are responsible for determining which control is used with which risk. These bodies also determine what the probability of a risk occurring is, as well as the severity of such an occurrence. The network is undirected, as unidirectional relationships between entities such as risks and risk owners do not seem realistic.

The network data were processed using the software suite ORA-Lite [24], while the SOMs were generated using the Viscovery SOMine Suite [20]. The data for the SOM consists of the 5 SNA metrics for each node. In total, 695 nodes are contained in the network, and a total number of 1738 links exist between them. The focus of the network is on managing real-world risks, and it was subsequently processed in a risk centric way, i.e. the relationships between the nodes are based on similar relationships to particular risks. This means that the relationship between a risk coordinator and a control owner, for example, is described only in terms of their shared relationship to the same risk.

With networks of this size, the large number of metric values that are produced can be quite complex. In order to help evaluate such complex data in a simpler, graphical way, the SOM algorithm is used. The SOMs can be used to quickly identify problem areas, which should make it easier to evaluate the data. Additionally, as SOMs are graphical in nature, they can be applied iteratively to investigate how the risk in a network changes over time, or as certain controls are introduced that aim to manage those risks.

When applying this technique to training risk managers, one of the aims is to highlight certain trends or groups that may pose a natural risk in the network. In doing so, the risk manager should be better informed of the nuances of the network and may be able to introduce more effective risk mitigation measures than would have been possible otherwise. Consider, for example, a network, such as the one shown in Fig. 1, with risks, control coordinators, and controls. If a SOM is developed for the network, the manager should be able to readily identify groups of nodes that have similar risk profiles based on their clustering. If a certain grouping of controls and control coordinators are found in the same cluster, for instance, it may indicate that there is a problem with the way in which the controls are managed. Alternatively, if all of the

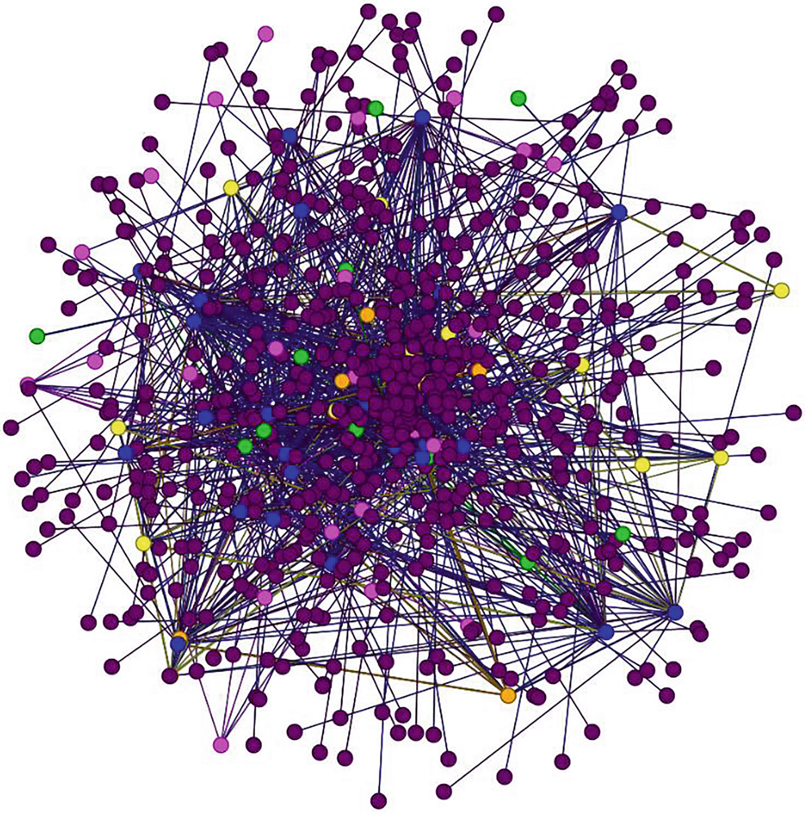

Fig. 1. Social network showing relationships between risks (blue), risk controls (purple), risk coordinators (yellow), risk owners (orange), control owners (pink) and the governing bodies (green) responsible for appointing the various role-players. (Color figure online)

control coordinators are grouped into one high risk cluster, it may be appropriate to introduce measures, such as policies, to address the risks posed by these nodes.

Another way in which the technique can be used in training is to monitor how the risk profile of certain clusters change when new controls are implemented to address the identified risks. As a SOM is graphical in nature, and the geographical structure of the map changes as the risk values for the nodes change, it should be possible to identify the changes in the network graphically. This is especially true for clusters that lose nodes, as the area of the map that the cluster occupies should be reduced. By using the SOMs as a graphical aid, the manager should be able to identify which approaches work best, and under which circumstances.

4 Results and Discussion

The SOM algorithm produced a map with three regions, or clusters, when applied to the network data. This map is shown in Figs. 2, 3, 4, 5 and 6. Each figure shows the same map, but with different colourations. The colourations are used to show how the values of the five measures differ for various nodes. The clusters are the same in each image, as one map, built using all five SNA metrics as node attributes, was obtained.

In Fig. 2, where the boundary spanner measure is used to colour the map, the red colouration that covers most of Cluster C2 indicates that the nodes in C2 pose a significant possible risk, as the nodes in this cluster have a much higher boundary spanner value than the nodes in other clusters. A cursory evaluation of the cluster's data shows that C2 exclusively contains all of the nodes that represent the risks. It should be noted that, while the risks are all found in the same cluster for this network, this will not necessarily be the case for all networks. As the boundary spanner metric indicates that a node's removal will completely isolate a part of the network, this suggests that the current structure in the network includes nodes that will be isolated if any of the risks are resolved. The network data itself shows that these nodes are primarily controls. If, for example, the risk "Corporate Brand" is completely resolved, i.e. if the company finds itself in a situation where there is no risk at all to the company brand, then the controls that exist to manage that risk, such as "Social Media strategy and protocols" and "Expert communications resources", are no longer needed. In order to ensure that these controls are not kept in place unnecessarily, additional measures need to be implemented.

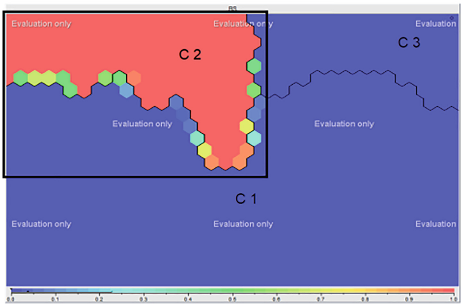

Fig. 2. SOM (boundary spanner) (Color figure online)

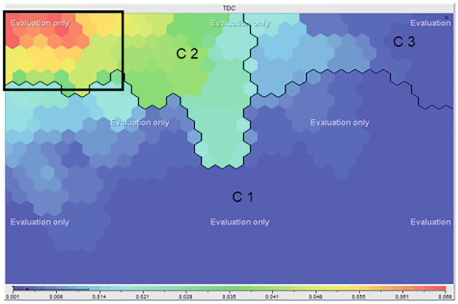

Fig. 3. SOM (total degree centrality) (Color figure online)

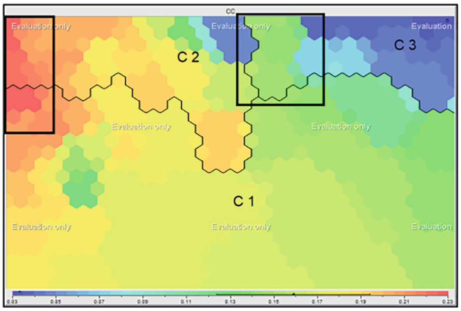

Fig. 4. SOM (closeness centrality) (Color figure online)

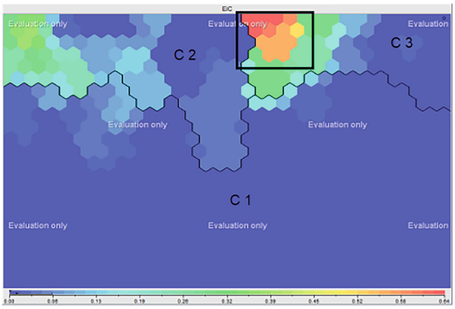

Fig. 5. SOM (eigenvector centrality) (Color figure online)

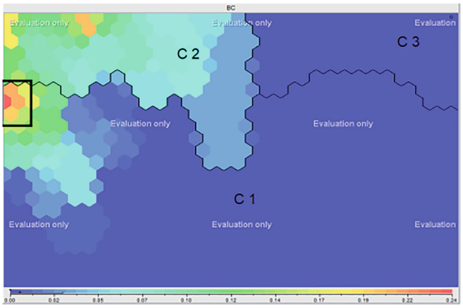

Fig. 6. SOM (betweenness centrality) (Color figure online)

The colouration, based on the total degree centrality, used on the SOM in Fig. 3 shows that there is an area in cluster C2 where the nodes have unusually high total degree centrality values, which is associated with a higher level of risk to integrity. Of the nodes in C2, there are six nodes, indicated by the red area in C2, that have a significantly higher amount of total degree centrality than the rest of the nodes. These nodes are the risks "Forest fires", "Environmental impacts", "Interruption to supply networks", "Waste Treatment Capacity", "Urban Resource Capacity", and "Rural Resource Capacity". All of these risks can have a significant impact if realised, which is why the influence they have is so substantial. This also means that any errors with regards to these risks, such as the risk of forest fires being over- or underemphasised, can have a significant impact on the information that is ultimately used in the network to manage other risks. If, for example, the integrity of the information with regards to the chances of a forest fire occurring is compromised, then there may not be enough water available to address the fire. If the fire affects any industrial assets, this may have environmental impacts, which in turn could negatively affect the company's corporate brand image. To protect the integrity of the information of these nodes, additional controls should be implemented.

An area with values much higher than the surrounding areas is shown to be present in cluster C3 in both Figs. 4 and 5. This area, which is situated on the left hand side of cluster C3 where it borders Cluster C2, is coloured green in Fig. 4 and green, orange, and red in Fig. 5. Additionally, a hotspot is present on the border between C1 and C2, as shown in Fig. 4. While the nodes in the hotspot area between C1 and C2 certainly pose a risk to confidentiality, the section of C3, where the nodes have much higher values for closeness centrality and eigenvector centrality than the rest of the nodes in the cluster, warrants further investigation. The higher closeness centrality measure of these nodes, shown in Fig. 4, suggest that they are a risk to confidentiality, whereas the very high eigenvector centrality measure shown in Fig. 5 indicate that they are an overall risk. There are two nodes in particular that fall into this region, one being a risk owner that is responsible for 11 out of the 26 risks, and the other being a risk coordinator that is responsible for 6 risks. The remainder of C3 is low risk and, as both of these nodes are in the cluster, it intimates that the risk posed by this risk owner and risk coordinator could be alleviated by reducing the number of risks that they are responsible for. Some of the risks could be transferred to other risk owners and coordinators. Alternatively, the risks can be co-owned and co-coordinated with owners and coordinators that are responsible for a smaller number of risks.

With the exception of the small area of low risk in C3, Fig. 4 shows that most of the nodes in the network have a high measure of closeness centrality. This suggests that controls should be in place to protect the confidentiality of the information in the network in general, as almost any one of the nodes could be responsible for compromising confidentiality.

Figure 6 highlights the existence of a single hotspot that exists in cluster C1 with regards to betweenness centrality. The hotspot, which is situated on the left hand side of C1 and has a red colour, contains a single risk coordinator, which poses a risk to both the integrity and the confidentiality of information in the network. In order to resolve this risk, the dependence on the specific risk coordinator should be reduced. The dependence of the risk coordinator can be reduced in one of at least two ways. The first method involves transferring some of the coordinator's responsibilities, such as risks, to another coordinator. This coordinator should preferably be located in cluster C1, as such a coordinator is likely to have a similar amount of power and influence. The second method that could be used is to employ an additional risk coordinator to take over some of the duties. This coordinator could also assume some of the duties of the risk coordinator in C3, thereby reducing the risk of two nodes simultaneously.

Based on all the SOMS shown in Figs. 2, 3, 4, 5 and 6, there are six risks, two risk coordinators, and a risk owner that pose potential risks to the overall security of information in this network. From this discussion, the advantage of using the SOM method to visualise SNA metric data is clear: by using SOMs to visualise SNA metric data, a relatively simple process can be followed in order to evaluate the risks in a network. The advantages are especially clear when compared to the process that would be needed in order to evaluate the risks in a network, such as the one shown in Fig. 1, or when using only raw data and statistical analysis. The use of SOMs in this manner gives managers the opportunity to evaluate risks graphically, as well as to compile risk discussion reports that do not require any prior knowledge of SOMs or SNA. This, in turn, could help improve the nature and quality of risk management, as a greater number of options and plans could develop as a result. Furthermore, SOMs allow for a way to systematically identify risks, and can also be used to monitor the progress and impact of risk mitigation strategies. Depending on the situation, it may be possible to identify positive or negative changes in the network almost instantaneously using this method. The central premise, i.e. that SOMs can be used to visualise SNA risk data, and in turn help educate managers of risks in their organisation in a quick and simple way, is therefore feasible.

5 Conclusion

Information security risk management is one of the most crucial parts of information security, but it is often complicated by the complexity of most organisational systems. In order to simplify the task of identifying risks in an organisation, SOM can be used to identify possible risks in an organisation by visualising SNA metric data of the organisation. A SOM, which clusters similar entities into geographically separate regions, is relatively simple to evaluate due to its graphical nature. When compared to other risk identification techniques that employ SNA metrics, which may require risk managers to process and evaluate large tables of numbers, the use of SOM may reduce the amount of work needed, as entities that pose a threat to the organisation can be identified with relative ease. Additionally, as a SOM is easier to evaluate, inexperienced risk managers may find the use of SOM less daunting than to use numerical data and statistical analysis. Finally, as a SOM provides an additional level of information that may not be readily apparent from the data, it could aid in educating risk managers of dangers in the organisation that may not be known, or obvious.

References

1. Wangen, G.: Information security risk assessment: a method comparison. Computer **50**(4), 52–61 (2017)
2. Armstrong, H., Armstrong, C., McCulloh, I.: A course applying network analysis to organizational risk in information security, In: South African Information Security Multi-conference pp. 204–214 (2010)
3. Dang-Pham, D., Pittayachawan, S., Bruno, V.: Investigation into the formation of information security influence: network analysis of an emerging organisation. Comput. Secur. **70**, 111–123 (2017)
4. Serfontein, R., Drevin, L., Kruger, H.: The feasibility of raising information security awareness in an academic environment using SNA. In: Drevin, L., Theocharidou, M. (eds.) WISE 2018. IAICT, vol. 531, pp. 69–80. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99734-6_6
5. Scott, J., Carrington, P.J.: The SAGE Handbook of Social Network Analysis. SAGE Publications, Thousand Oaks (2011)
6. Tsui, E., Liebowitz, J.: Linking social network analysis with the analytic hierarchy process for knowledge mapping in organizations. J. Knowl. Manag. **9**(1), 76–86 (2005)

7. Dang-Pham, D., Pittayachawan, S., Bruno, V.: Applying network analysis to investigate interpersonal influence of information security behaviours in the workplace. Inf. Manag. **54** (5), 625–637 (2017)
8. Boulet, R., Jouve, B., Rossi, F., Villa, N.: Batch kernel SOM and related Laplacian methods for social network analysis. Neurocomputing **71**(7), 1257–1273 (2008)
9. Kohonen, T.: The self-organizing map. Neurocomputing **21**(1–3), 1–6 (1998)
10. De la Hoz, E., De la Hoz, E., Ortiz, A., Ortega, J., Prieto, B.: PCA filtering and probabilistic SOM for network intrusion detection. Neurocomputing **164**(Suppl. C), 71–81 (2015)
11. Hunt, R., Hill, S.: Using security logs to identify and manage user behaviour to enhance information security. In: 14th European Conference on Cyber Warfare and Security, p. 111. Academic Conferences Limited (2015)
12. López, A.U., et al.: Analysis of computer user behavior, security incidents and fraud using self-organizing maps. Comput. Secur. **83**, 38–51 (2019)
13. Bäck, T., Kok, J.N., Rozenberg, G.: Handbook of Natural Computing. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-540-92910-9
14. Pal, C., Hirayama, S., Narahari, S., Jeyabharath, M., Prakash, G., Kulothungan, V.: An insight of world health organization (WHO) accident database by cluster analysis with self-organizing map (SOM). Traffic Inj. Prev. **19**(sup1), S15–S20 (2018)
15. Nakayama, H., et al.: Comparative transcriptomics with self-organizing map reveals cryptic photosynthetic differences between two accessions of north american lake cress. Sci. Rep. **8** (1), 3302 (2018)
16. Gu, F., Cheung, Y.-M.: Self-organizing map-based weight design for decomposition-based many-objective evolutionary algorithm. IEEE Trans. Evol. Comput. **22**(2), 211–225 (2018)
17. Kuo, R.J., Rizki, M., Zulvia, F.E., Khasanah, A.U.: Integration of growing self-organizing map and bee colony optimization algorithm for part clustering. Comput. & Ind. Eng. **120**, 251–265 (2018)
18. Lee, Y.: Using self-organizing map and clustering to investigate problem-solving patterns in the massive open online course: an exploratory study. J. Educ. Comput. Res. (2018). https://doi.org/10.1177/0735633117753364
19. Fausett, L.V.: Fundamentals of Neural Networks: Architectures, Algorithms, and Applications. Prentice-Hall, Englewood Cliffs (1994)
20. Viscovey SOMine. www.viscovery.net/somine. Accessed 10 Feb 2019
21. Au, C.H., Fung, W.S., Tses, A.: An investigation on the relationship between control self-assessment, cloud security, and cloud-related business performance-using partial least squares, In: Industrial Engineering and Engineering Management (IEEM), pp. 1879–1883. IEEE (2016)
22. Armstrong, H., McCulloh, I.: Organizational risk using network analysis, In: South African Information Security Multi-conference, pp. 132–141 (2010)
23. Hanneman, R.A., Riddle, M.: Introduction to Social Network Methods. University of California, Oakland (2005)
24. ORA-Lite. www.casos.cs.cmu.edu/projects/ora. Accessed 24 Apr 2018
25. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 2nd edn. The MIT Press, Cambridge (2001)
26. Borgatti, S.P.: Centrality and network flow. Soc. Netw. **27**, 55–71 (2005)