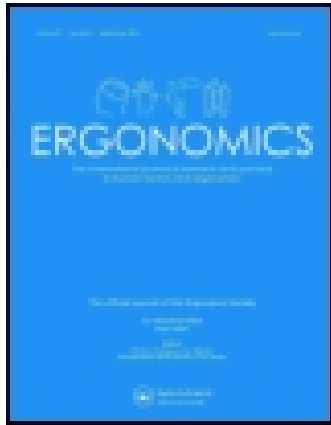


This article was downloaded by: [Michigan State University]

On: 26 February 2015, At: 09:56

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Ergonomics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/terg20>

From trees to forest: relational complexity network and workload of air traffic controllers

Jingyu Zhang^a, Jiazhong Yang^b & Changxu Wu^c

^a Institute of Psychology, Chinese Academy of Sciences, Beijing, China

^b Flight Technology College, Civil Aviation Flight University of China, Guanghan, Sichuan, China

^c Department of Industrial & Systems Engineering, State University of New York at Buffalo, Buffalo, NY, USA

Accepted author version posted online: 13 Feb 2015. Published online: 24 Feb 2015.



[Click for updates](#)

To cite this article: Jingyu Zhang, Jiazhong Yang & Changxu Wu (2015): From trees to forest: relational complexity network and workload of air traffic controllers, *Ergonomics*, DOI: [10.1080/00140139.2015.1009498](https://doi.org/10.1080/00140139.2015.1009498)

To link to this article: <http://dx.doi.org/10.1080/00140139.2015.1009498>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

From trees to forest: relational complexity network and workload of air traffic controllers

Jingyu Zhang^{a*}, Jiazhong Yang^{b1} and Changxu Wu^{c*}

^a*Institute of Psychology, Chinese Academy of Sciences, Beijing, China;* ^b*Flight Technology College, Civil Aviation Flight University of China, Guanghan, Sichuan, China;* ^c*Department of Industrial & Systems Engineering, State University of New York at Buffalo, Buffalo, NY, USA*

(Received 9 April 2014; accepted 14 January 2015)

In this paper, we propose a relational complexity (RC) network framework based on RC metric and network theory to model controllers' workload in conflict detection and resolution. We suggest that, at the sector level, air traffic showing a centralised network pattern can provide cognitive benefits in visual search and resolution decision which will in turn result in lower workload. We found that the network centralisation index can account for more variance in predicting perceived workload and task completion time in both a static conflict detection task (Study 1) and a dynamic one (Study 2) in addition to other aircraft-level and pair-level factors. This finding suggests that linear combination of aircraft-level or dyad-level information may not be adequate and the global-pattern-based index is necessary. Theoretical and practical implications of using this framework to improve future workload modelling and management are discussed.

Practitioner Summary: We propose a RC network framework to model the workload of air traffic controllers. The effect of network centralisation was examined in both a static conflict detection task and a dynamic one. Network centralisation was predictive of perceived workload and task completion time over and above other control variables.

Keywords: relational complexity network; air traffic control; workload; network centralisation

1. Introduction

In 2013, thirty-three million aircraft carrying more than three billion passengers flew around the world. The safety and efficiency of these flights are guaranteed by ground-based air traffic controllers (ATCos) who monitor and organise the flow of air traffic and direct aircraft in certain situations to prevent collisions. Although the current air traffic management (ATM) system has been running effectively for more than half a century, it is now reaching its maximum capacity. According to a recent forecast made by the International Civil Aviation Organization, passenger and freight air traffic will double by 2030 (ICAO 2013). As airspace is limited, if nothing is done, air traffic complexity and controllers' workload will become unmanageable. In response to this rapidly growing demand, more effective usage of airspace, while maintaining current safety levels, is needed. As it has long been recognised that the workload of ATCos is an important constraining factor on system capacity and safety, a precise understanding and modelling of controllers' workload is vital for the development of future ATM systems such as NextGen or SESAR.

Researchers have attempted to establish a link between air traffic complexity and the cognitive workload of controllers for decades (for recent reviews see Athènes et al. 2002; Hilburn 2004; Loft et al. 2007; Prandini et al. 2011). Early attempts focused on producing a task complexity indicator as a proximate for controllers' mental workload by aggregating aircraft-level information in a sector: for example, total number of flights, number of climbing/descending aircraft and so on (Mogford et al. 1995; Laudeman et al. 1998; Gianazza 2010). However, this approach has several drawbacks: first, the large number of factors and highly variable coefficients between sectors make it difficult to use in practice; second, the data-driven nature makes it less likely to shed light on the inner cognitive processes of ATCos. Therefore, recent studies have paid more attention to the cognitive processes and have suggested that the relational information between aircraft dyads is a more important workload driver (Boag et al. 2006; Loft et al. 2009; Rantanen and Nunes 2005; Stankovic, Raufaste, and Averty 2008). While these studies have greatly improved our ability to quantify the cognitive complexity in ATCos' decision-making, it has not yet been established how to integrate the pair-level information into a higher holistic level. From a systems theory perspective (Anderson 1972; Neal et al. 2014; Simon 1991; Walker et al. 2010), complexity is an emergent property of the interactions amongst the constituent elements which is not merely the sum of its parts. On the one hand, the growing interest in the heuristics and strategies used in ATCos' decision processes has suggested that the linear combination of low level data may not be enough and a certain overall pattern might be more important but difficult to quantify (Loft et al. 2007; Neal and Kwantes 2009; Vuckovic, Kwantes, and Neal 2011). On the other hand, although

*Corresponding authors. Email: zhangjingyu@psych.ac.cn; changxu.beijing@gmail.com

complexity theory has been introduced in ergonomics research (for a review, see Walker et al. 2010), research to quantify the cognitive complexity of ATCos is not well established (but see Delahaye and Puechmorel (2000, 2010), for a dynamic system approach without fully considering human factors).

Analysis of the topological structure of the cognitively processed physical relationships among all aircraft dyads in a given sector might provide a new way to understand this issue. By borrowing concepts from network theory and relational complexity (RC) metric, we have attempted to represent the relational information between aircraft dyads in a network as well as investigate the effects of some global network properties on ATCos' mental workload associated with conflict detection and resolution. In the following sections, we first review the basic task requirements of ATCos and the ongoing efforts in ATC workload modelling. Second, we introduce a network-based framework and provide reasons why the properties of such networks (size and structure) can influence controller's workload. Next, we propose a general computational indicator representing the network structure. Finally, we provide experimental evidence showing that this indicator that can be used as a predictor of ATCos' workload in addition to other controlling variables.

1.1. How controllers process the relational information of aircraft dyads

The primary goal of en-route ATCos is to ensure that the applicable separation minima between an aircraft and a hazard (other aircraft, terrain, weather, etc.) are never violated. The most commonly used separation minima are 5 nautical miles for lateral separation and 1000 feet for vertical separation. In fulfilling this goal, ATCos can use both strategic and tactical measures to avoid collision. While the former can be realised through pre-departure or early route planning, the tactical measures are of primary concern in our research which consists of conflict detection and resolution. In using such measures, ATCos have to monitor the flight information (speed, altitude and direction), predict their future trajectories and assess the possibility of future conflicts. When they find any potential danger, they need to formulate a solution, implement it and monitor its execution in an iterative manner (ICAO 2005).

Given the very nature of this task (detecting and resolving conflicts among aircraft pairs), it is not unexpected that the relational information among aircraft plays an important role in controller's perception, decision-making and actions. Some relational characteristics directly contribute to the cognitive complexity of ATC tasks. For example, Neal and Kwantes (2009) have pointed out that conflict detection accuracy and response time vary as a function of angle between two aircraft's trajectories. Also, it is more difficult for controllers to handle aircraft pairs when their minimum distance, the distance between two aircraft when they are closest to each other, is lower or around the separation standard (Loft et al. 2009; Neal and Kwantes 2009; Neal et al. 2014). In addition, if relational information is provided in an easy-to-understand manner (e.g. visualised relative position vectors), controllers' performance can be significantly improved (Vuckovic et al. 2013).

Beyond angle and distance, Boag et al. (2006) suggested that another important predictor of workload is RC which is determined by the number of different parameters (speed, altitude, headings, etc.) a controller needs to consider simultaneously in any dyad-based conflict detection task. For example, RC is very low when two level flight aircraft are on separate levels or two same level aircraft are in parallel or diverging directions ($RC = 0$). In these cases, the cognitive effort for ATCos to ensure no potential conflict emerges is relatively small (also see Rantanen and Nunes 2005). Meanwhile, a medium RC arises when two flights fly along converging paths at the same level ($RC = 1$) or one aircraft crossing levels ($RC = 2$). In these cases, ATCos have to assess the trajectories of both flights and see whether they converge at a certain time point. An even higher RC may occur when two aircraft with different speed are flying into or out of the same path, with one aircraft climbing/descending across the level of another ($RC = 3$ or 4). ATCos in these cases have to figure out the trajectories using all relevant variables (speed, direction, climbing/descending rate) in order to guarantee both lateral and vertical separation. In their research using 60 static ATC scenarios composed, in each case, of three aircraft, Boag et al. (2006) found that linearly aggregated RC values (from 0 to 12) can predict additional variance of controller's workload and reaction time over and above aircraft count and other relational information (e.g. averaged minimum distance of all aircraft pairs in each scenario).

While these findings are very promising, there is still some room for improvement. First, most research on relational information has only investigated situations with a relatively small number of aircraft [e.g. two aircraft in Loft et al. (2009), three in Boag et al. (2006), and four in Vuckovic et al. (2013)]. ATCos generally need to deal with 12–20 aircraft in a sector, so it is important to understand how pair-based findings could be aggregated into the sector level as the number of aircraft increases. While some researchers have suggested a linear combination might be enough (Boag et al. 2006), more recent findings suggest a nonlinear relationship might exist as ATCos can manage their workload dynamically (Loft et al. 2009, 2007). More importantly, it is necessary to know whether there are any higher-order constructs representing the relational patterns that cannot be produced by merely analysing the basic constituents. Although it has been documented repeatedly in expert problem solving research that experienced task performers are more likely to recognise the holistic pattern of a situation rather than adopting a step-by-step serial process in dealing with their tasks (Gobet 1997; Kahneman

and Klein 2009), in the domain of ATC research, very few studies have paid attention to the possible holistic patterns of aircraft interactions.

It is worth noting that Delahaye and Puechmorel (2000, 2010) have adopted a dynamic system approach. Based on position and speed vectors of each aircraft in the given sector, the order-ness and predictability of the whole system can be analysed. Although this approach offers very important insights from the system perspective, it does not take human factors into consideration (Delahaye and Puechmorel 2010). Therefore, the estimated system complexity may not necessarily correspond to cognitive complexity and mental workload of human operators. This mismatch can occur for different reasons. In the first place, operators are confined by the human-machine interface (2D radar screen) and limited cognitive capacity but the system approach always has a 'god-like view' that all aircraft information (in three-dimensional Euclidean space plus the time dimension) is available for thorough calculation. In the second place, although the dynamic system approach has taken the spatial distribution of complexity into consideration, it does not fully consider the distribution of the cognitively processed physical relationships around different aircraft (which is independent of real distance). In the next paragraphs, we will show how this distribution is important for controllers' decision-making.

1.2 A network representation of aircraft relational information

We adopt here a new approach to analyse the holistic patterns of ATC tasks. We first created a RC network (RCN) based on RC metrics developed by Boag et al. (2006) to account for the pair-wise relationships and combined all these relations to build a network based on a graph theory approach. In graph theory, a network is defined as a set of nodes together with a set of pair-wise relationships (edges) among them. In the RCN, each node represents a single aircraft, whereas each edge represents the degree of RC in detecting and resolving conflict between aircraft dyads. This reflects the level of difficulty deciding whether two aircraft might have a conflict and how to cope with it. In the study by Boag et al. (2006), RC is an ordinal variable varying from 0 to 4 without direction (higher values indicate higher complexity). In our research, we downgraded this into a dichotomous variable such that each edge is either 1, representing highly complex relations, or 0, representing low complexity (for details about the operationalisation, see the Method section).

When all values of dyad RC (edges) among all aircraft (nodes) are identified, an adjacent matrix can be built and a network can be established and represented in a graph. Several network properties can be given accordingly. In the first place, network size is equal to the number of nodes (N). Second, total RC is equal to the number of all existent edges (RC, in the dichotomous network). Further, we can identify two simple but representative network patterns that are characterised by the different distribution of degrees (i.e. number of connections) of each node. The first star-like network is highly centralised in which one central node is linked to all other nodes among which no connections are present. In contrast, in a chain-like network, there is no highly centralised aircraft and all nodes have similar numbers of connections.

To make the framework clear, we can provide an example to illustrate how these networks are established and represented. Figure 1 shows two scenarios: A and B. In scenario A, while all aircraft have crossing trajectories horizontally, aircraft A is climbing through, or to, the levels of all other aircraft (A2, A3 and A4) at the same time. According to previous studies (Boag et al. 2006; Rantanen and Nunes 2005), flights that are clearly vertically separated cause low RC but flights that are climbing/descending through each other's routes can cause high RC. Thus, we code the relationships A1-A2, A1-A3 and A1-A4 as RC = 1 and all other relationships (A2-A3, A2-A4, A3-A4) as RC = 0. Using the same principle, we can also code the relationships in scenario B. In this case, B1 is climbing through the level of B2, but is vertically separated from B3 and B4. Meanwhile, B2 is further climbing through the level of B3, but not B4, whereas B3 is climbing through the level of B4. Therefore, the RC of the relationships B1-B2, B2-B3, B3-B4 can be coded as RC = 1 and all other relationships (B1-B3, B1-B4, B2-B4) as RC = 0. Based on this type of coding, we can produce an adjacent matrix and a graphic representation of this network (as shown below in Figure 1) where aircraft and flight information are replaced by nodes and edges.

Clearly, the two scenarios have almost identical patterns in terms of number of aircraft (network size), number of aircraft changing altitude, merging patterns and total RC (number of edges), but they are different in the network patterns. Network A is a typical star-like network which has a very centralised node A1, while network B is a typical chain-like network. From this, we can begin to analyse how these patterns can cause differences in controllers' cognitive processes and mental workload.

1.3 Network properties and controllers' workload

Having established the network framework, we can now discuss how to use it to analyse controllers' decision-making and mental workload. In the first place, studies have demonstrated that controllers' workload increases as the number of aircraft increases (Boag et al. 2006; Loft et al. 2007; Mogford et al. 1995). The new terminology in our study (network size) does

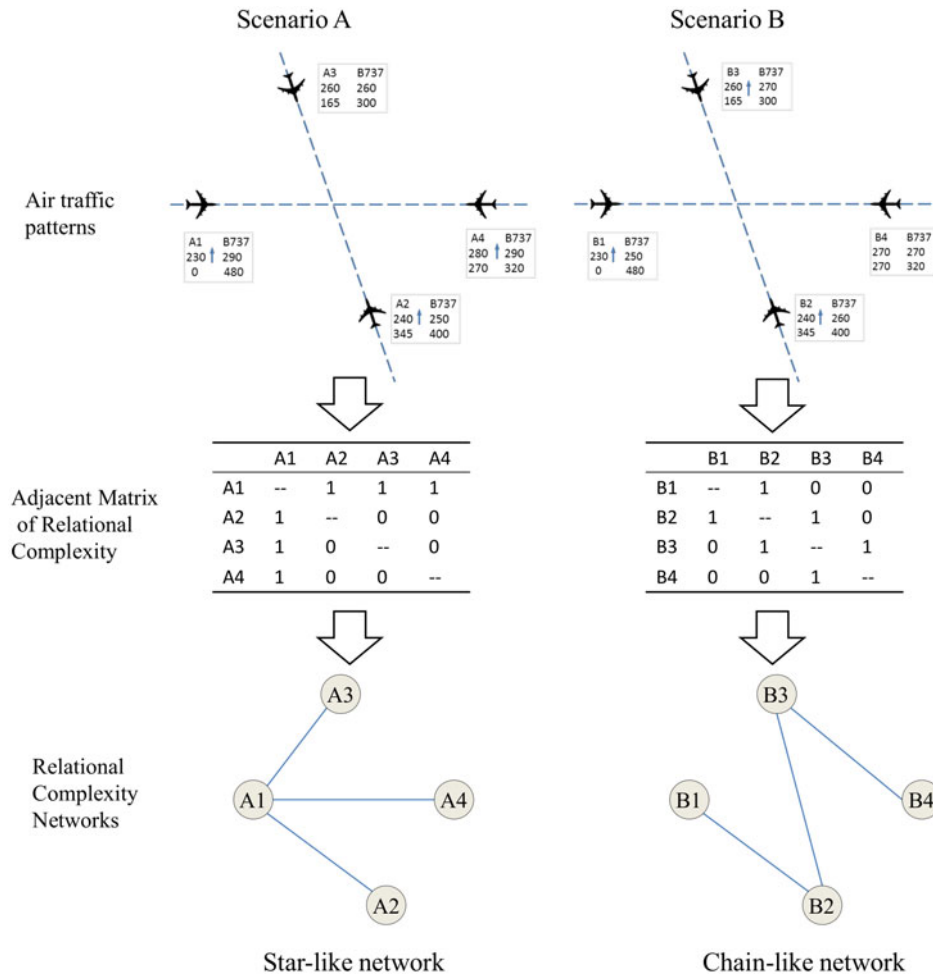


Figure 1. The establishment of relational complexity network from air traffic patterns.

Note: The flight information:

Row 1: call sign and plane type;

Row 2: current and destined altitude (in hundred feet), with arrows indicating whether descending ↓ or ascending ↑;

Row 3: flight direction (in degrees) and speed (in miles per hour).

not change the influence of aircraft count in producing ATC workload. Nevertheless, it is a useful test to see whether the experiment conducted to examine our framework can successfully manipulate task difficulty in the first place. Therefore, we propose our first hypothesis:

H1: controllers' mental workload and reaction times increase as number of flights increases.

Second, total RC (number of edges in our network) also greatly influences controllers' workload. Given the same number of aircraft, the overall complexity can be quite low when all planes are vertically separated. On the other hand, it can be very high, when all aircraft are crossing each other's routes both horizontally and vertically, etc. As already noted, while this argument is theoretically sound, it has only been empirically tested with a small number of flights (three aircraft in Boag et al. 2006). It is not yet known whether this metric is still effective in situations where more aircraft are involved, and when dichotomous coding is used. Therefore, in our research, the second aim is to test whether controllers' workload increases as total RC increases.

H2: controllers' mental workload and reaction times increase as total RC values increases.

Based on these premises, we propose that the cognitive responses of ATCos may be different when facing different network patterns. While many network properties can be utilised, our research will focus on the centralisation of a network. More specifically, we propose that a centralised network (such as a star-like network) will cause a lower workload and faster

response time for both conflict detection and resolution. The reason is related to two possible mechanisms: first, when facing a star-like pattern, controllers' attention is likely to be more focused on a specific node (the central one) because during their visual search, they can observe that all relations to be processed have one thing in common, namely, they are connected to the central node. Therefore, in their subsequent mental calculation of all possible potential conflicts, they can use the trajectory of the central aircraft in their visual working memory as a constant reference or template and compare it with the trajectories of all other aircraft. This 'no-need-to-change-reference' can be very advantageous in the conflict detection process for although human visual working memory can store different items simultaneously, it can only use one active reference or template at a time (Olivers et al. 2011). For example, when controllers are facing a situation as shown in Figure 1(a), they need to consider whether A2, A3 and A4 would have conflict with A1. In doing so, they do not need to update the data for A1, but when controllers need to consider situations as shown in Figure 1(b), they have to change the reference aircraft at least twice (e.g. B2 and B3).

Second, in the conflict resolution process, controllers can prevent all potential conflicts in future by actively modifying the trajectory of the central node in a star-like network, but this is not easy when facing a decentralised network. For example, in Figure 1(a), if controllers change the cleared level of A1 to 23,000 feet, all potential conflicts can be resolved by this 'once-and-for-all' shortcut. However, to resolve the potential conflicts in the case of Figure 1(b), controllers have to change the properties of at least two aircraft (e.g. B2 and B3). It is worth noting that in some particular situations, when the central node cannot be easily found or controlled, the controllers may have an even higher workload, as they have to intervene with all other flights that are related to the focal one. For instance, the situation in which a military aircraft traverses many different aircraft routes can be seen as a typical star-like network. However, as controllers often do not have the right to fully control this kind of aircraft, they have to change the routes of all other related airplanes and thus encounter a higher workload. However, while the new framework presented in our research can explain the high workload of this abnormal event effectively, we restrict our discussion here to more commonly encountered situations.

1.4 Centralisation: a computational indicator of the network structure

Following the above analyses, a star-like network can result in a lower mental workload and faster response times among controllers in both conflict detection and resolution processes. However, a pure star-like network is difficult to find in real situations and this categorisation can only provide a dichotomous variable (either star-like or not). To make this framework practical, a more general and mathematical quantification of the 'similarity to a star-like network' is needed. Network centralisation is such a measure. Introduced by Freeman (1978), centralisation (C) describes how the nodes in a network are organised around a focal point. A network with higher centralisation index is more similar to a star-like network. Centralisation is calculated by the following formula:

$$C = \frac{\sum_{i=1}^n (C_{\max} - C_i)}{\max[\sum_{i=1}^n (C_{\max} - C_i)]} \quad (1)$$

In formula (1), C_i refers to the centrality degree (the number of linked edges for any given node i) and C_{\max} is the greatest centrality degree among all nodes in a given network. Therefore, the numerator evaluates the total difference in the number of connected edges between the most central node and all other nodes. The denominator is the largest possible total difference in an n -node network. In fact, the denominator reaches its maximum in an n -node star-network of $n - 1$ edges in which a central node is linked with all other $n - 1$ nodes. Therefore, the $C_{\max} = n - 1$, and $\sum_{i=1}^n (C_{\max} - C_i) = ((n - 1) - (n - 1)) + (n - 1 - 1) \times (n - 1) = n^2 - 3n + 2$. Using formula (1), the network centralisation of Figure 1(a) and (b) is $((3 - 1) \times 3 + (3 - 3) \times 1) / (4^2 - 3 \times 4 + 2) = 100\%$ and $((2 - 1) \times 2 + (2 - 2) \times 2) / (4^2 - 3 \times 4 + 2) = 33.3\%$, respectively. Indeed, many network analysis software packages (e.g. Ucinet, see Borgatti, Everett, and Freeman 2002) can do this computation conveniently. So, centralisation can be easily computed and used to represent different RC networks with a universal scale. In our study, we explore through empirical experiments the following hypothesis:

- H3: controllers' mental workload and reaction times decreases as the centralisation value of the RC network increases (when the network is more similar to a star-like network).

1.5 Research design and control variables

So far, we have delineated the network perspective to analyse the workload of air traffic control tasks. Established on RC metrics, this framework can offer a holistic approach and provide new insights that have not been addressed elsewhere.

Network-related properties seem to be more relevant for human heuristic processing in expert decision-making than pair-wise analysis, but to the best of our knowledge, no such models have been proposed in the domain of ATC workload research. In order to provide an empirical test of this framework, we conducted two experiments to test the hypotheses we have proposed.

As we have argued before, the benefit of a centralised network in reducing workload might result from two different mechanisms: (1) it provides a ‘no-need-to-change-reference’ benefit in visual searches and (2) it provides a ‘once-and-for-all’ shortcut in making interventions. Differentiating the two mechanisms is important but difficult as they are highly intertwined with each other in real task conditions. Therefore, we combined both static and dynamic task designs to test the aforementioned hypotheses. In Study 1, we conducted a static ATC task similar to the study by Boag et al. (2006) but using scenarios with more aircraft and varied network patterns. In Study 1, the participants were asked to detect all potential conflicts first and only after that were they able to resolve them. This study first serves as a direct expansion on Boag et al. (2006) to show the incremental predictive power of the network characteristic. Moreover, as the conflict resolution strategies are not very relevant in pure static situations, any findings of Study 1 can be largely ascribed to the effect occurred in the detection process. In Study 2, we used dynamic tasks similar to real situations in which while the participants also had to report the conflicts, they could resolve these conflicts at any time during the ongoing scenario. In this way, they can prevent potential conflict without adopting an exhaustive visual search process. If we still observe the effect of network structure on workload and task completion time in this task setting, it provides evidence for the ‘once-and-for-all’ benefit in the resolution process.

In both tasks, the three hypotheses (H1–H3) were tested by fully considering several important control variables that have been identified in previous studies. The first three were used in the original study by Boag et al. (2006) testing the effect of RC metrics in a three-aircraft situation, while the fourth variable was used by Neal et al. (2014) in their more dynamic tasks to test a multilevel model of ATCos’ workload.

- (a) Weighted distance (WD) of minimum separation. This variable describes the minimum possible distance in the three dimensions between two aircraft weighted by minimum vertical and horizontal separation standards:

$$WD = \sqrt{\min_{t \geq 0} \left(\frac{\text{lateral separation at } t_i \text{ (nm)}}{5 \text{ nm}} \right)^2 + \left(\frac{\text{vertical separation at } t_i \text{ (ft)}}{1000 \text{ ft}} \right)^2}.$$

In other words, this variable is measured by a separation unit. It should be noted that, using the Euclidean distance without the weighting would underestimate the effect of vertical separation on controllers’ conflict detection (because it is generally smaller in scale than horizontal separation) (Boag et al. 2006; Vuckovic et al., 2013, etc.). By averaging all pair-wise values in each scenario, previous research has found the scenario-based WD was negatively related to controllers’ workload (Boag et al. 2006). In the current research, we predicted that the averaged value of this variable in a scenario would be negatively associated with controllers’ workload rating and task completion time.

- (b) Conflict duration (CD). This refers to how long two aircraft would be in violation of the separation standard (in the case of pairs that have potential conflicts). A longer duration indicates that the conflict is easier to detect, causing lower workload. In our research, we also postulated that scenario-based CD (the mean of all pair-wise values in each scenario) would be negatively related to workload ratings and task completion time.
- (c) Time to conflict (TC). Termed by Boag et al. (2006) as temporal proximity, this is the amount of time that remains until an aircraft pair violates the separation standard (for a conflicting pair) or reaches their distance of minimum separation (for a non-conflicting pair). Although the researchers believed that a smaller TC should produce higher time pressure and higher workload, empirical evidence has been mixed (Boag et al. 2006; Remington et al. 2000). However, this might have resulted from their use of static conflict detection tasks in which the aircraft did not move, thus time pressure may not have caused any difference. As we combine both static and dynamic tasks in our research, we keep this variable (also averaged across all aircraft pairs) in our further analysis.
- (d) (iv) Violation Count. This variable is the number of aircraft pairs that would violate the minimum separation standard in a certain period of time given their current flight plans. It is hypothesised that more potential conflicts result in higher workload.

2. Study 1

Study 1 was conducted to investigate participants’ workload and task completion time in static conflict detection tasks (Boag et al. 2006). In doing these tasks, operators were shown a series of static air traffic scenarios and required to find all potential conflicts and mark them accordingly. In this way, controllers were prevented from managing their workload by

choosing particular strategies: e.g., making early interventions to avoid violations that they are not absolutely sure of or situations that required difficult calculations (preventive intervention). As a result, the difference in participants' workload and task completion time can largely reflect the task demands on visual search and mental calculations rather than the choice of intervention strategies.

2.1. Method

2.1.1. Participants

Thirty-seven trainee controllers were recruited from the air traffic control training centre at the Civil Aviation Flight University of China in return for course credits and 120 Yuan payment (20 US dollars, approximately). At the time of the experiment, these trainees had completed four years of college study in air traffic control management, one year simulator training and had also received the entree level air traffic control certificate. Due to an oversight in the pre-task training session, one participant failed to use the supportive tools (for details, see below) throughout the whole session, thus producing an extremely long task completion time (4 SDs above average). As a result, only the remaining 36 participants (33 males and 3 females) were used for final data analysis. They were aged between 21 and 27 ($M = 22.78$, $SD = 1.17$).

2.1.2. Conflict detection and resolution task

We followed the design used by Boag et al. (2006) for the static conflict detection task in which aircraft would not move during the whole scenario and no time limit was set. However, the participants were asked to treat this as real life task. We used an ATC-Simulator, a medium fidelity air traffic control simulation platform developed for this study, to conduct our research. All scenarios were presented in a 220 nm \times 180 nm en route sector. Each aircraft that appeared on the radar screen had an information block containing the call sign, the aircraft type, the current and target flight level (in hundred feet), altitude change (\uparrow for climbing and \downarrow for descending), heading (in degrees) and ground speed (in nautical miles/hour).

At the beginning of each scenario, a certain number of aircraft appeared within the sector boundary. The participants were asked to detect all aircraft pairs that would violate the minimum separation standard. When they found such conflicts, they used the 'detect' option from a dropdown menu by clicking the icon of any flights that were in conflict and moved the mouse cursor to another aircraft icon that they found to be in potential conflict. In selecting the new aircraft icon, confirmation including the call signs of both aircraft would appear on the right side of the screen. By selecting the 'confirmation' option, the participants were able to mark the potential conflicting pair. When they marked all detected conflicts, they would click the 'Finish Detection' button. Total conflict detection time was measured in seconds from the start of each trial to the time when controllers pressed the final button.

Thereafter, they were asked to offer solutions for all the potential conflicts by changing the flying parameters (altitude, speed and direction) by pressing the corresponding figures within the flight information window on screen. In doing so, a dropdown list of options appeared from which participants were able to choose their instructions to the plane. For example, if they wanted an aircraft flying at 20,000 feet to climb to 25,000 feet, they would click the figure 200 in its information window, choose 250 from the dropdown list containing numbers ranging from 100 to 300 and click the 'uplink' option. After they had resolved all potential conflicts in this way, four questions were asked to quantify their perceived workload during the whole task. Responses were provided using an 8-point likert scale. The questions were: (a) 'How difficult was it to find all potential conflicts?' (1 = very easy, 8 = very difficult); (b) 'How difficult was it to resolve all the conflicts?' (1 = very easy, 8 = very difficult); (c) 'How much time pressure did you perceive during this task?' (1 = very low, 8 = very high); (d) 'To what extent did you think it is necessary to make a preventive intervention?' (1 = not very much, 8 = very much). The summed scores for all four questions were used as the measure of workload. We calculated the α coefficients of the four-item workload scale for all 18 scenarios. The minimum value among the 18 values was .73, the maximum was .91 and the average was .84.

Several supporting tools were provided for the controllers. First, a 20 nm \times 10 nm scale maker which could be moved to any location was provided initially at the bottom left of the screen. Second, a range bearing line function was provided which could be used to show the future trajectory of any aircraft given current velocity and direction (see green lines in Figure 2). Third, a distance and time calculation function was provided: when participants moved the mouse cursor to an airplane, they could hold down the left button on the mouse and move the cursor to any point on the screen, thus revealing information including directions, distance and time to reach the particular point, given its current velocity (see the red line and white words in the centre of Figure 2).

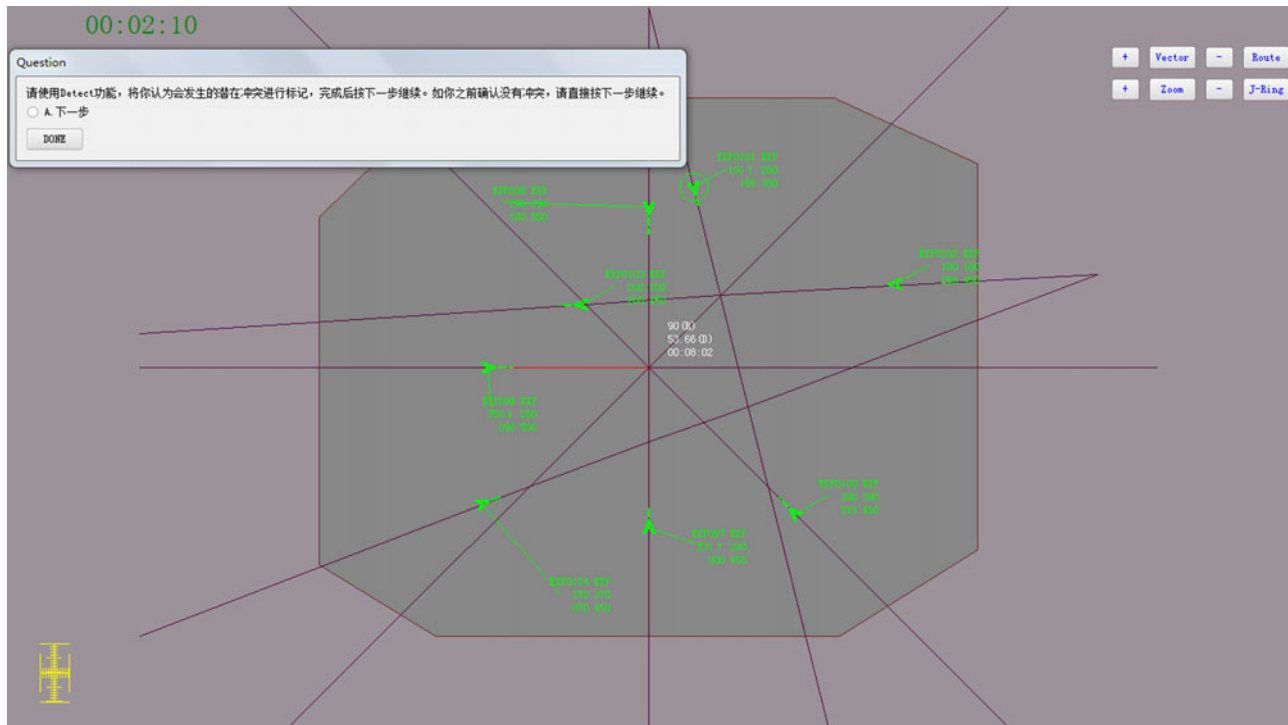


Figure 2. Simulator interface.

2.1.3. Scenarios

A pool of scenarios was first designed by subject matter experts. For each, the RC value of all aircraft pairs was evaluated by two independent raters using a dichotomised RC metric adapted from Boag et al. (2006). RC was coded as 0 if two planes had non-converging paths or had converging paths but were vertically separated from each other. All other more complex situations were coded as 1. Indeed, most of these situations were conditions involving climbing or descending aircraft with converging paths (the RC = 2 conditions in the original study of Boag et al. 2006). Complex situations with merging or diverging paths (the RC = 3 and RC = 4 conditions) were not used. Full agreement was reached for 18 scenarios which were later used as the formal experiment.² Next, following the procedure we described in section 1.2, the RC network was established and the network centralisation index was calculated using Ucinet 6.02 (Borgatti, Everett, and Freeman 2002). In total, there were 25 conflicts among all 312 potential interactions in the scenarios used (for details, see Appendix).

2.1.4. Procedure

Upon arrival, the participants were asked to sit by a computer with a 23-inch-wide LED monitor. They were then helped to familiarise themselves with the simulator, particularly how to use the supportive tools and the response options. The participants were told that the climbing/descending rate of aircraft were all set at 1000 feet per minute in all scenarios and completed three practice scenarios before commencing the formal task. During the task, each participant completed all 18 formal and other 6 filler scenarios in a random manner with a 2-minute break after every 8 scenarios.³ The overall duration of the whole experiment generally lasted for 2 hours.

2.3. Results

Before presenting the regression results, the general task performance was described. On average, participants reported 2.09 conflicts ($SD = .79$) and spent 99.41 seconds ($SD = 71.85$) completing each scenario. The final miss rate was 14.6%.⁴ The insignificant correlation between reaction time and miss rate ($r = -.27, ns.$) suggests that there was no accuracy–speed trade-off.

Given the data structure of the study (scenarios performed by different individuals), and our interest in the additive power of the three variables (flight numbers, total RC and network centralisation), we conducted a series of multilevel

analyses (Raudenbush and Bryk 2002) using the HLM 6.02 package on the two dependent variables: perceived workload and task completion time. The means, SDs and zero-order correlations of all variables were shown in Table 1.

To conduct the multilevel analysis, a null model with no predictors at both levels was first estimated to quantify the within-individual and between-individual components of variance for each dependent variable. The intra-class correlations (ICCs) indicated that 32.6% of the variance in the perceived workload, and 28.2% of the variance in task completion time, was at the between-individual level. The existence of such a large variance justified our usage of multilevel analysis. It should be noted that, in this paper, only level 1 predictors were adopted; the analysis of level 2 (individual) predictors would be discussed elsewhere.

Next, we tested our hypotheses in four hierarchical models in which the incremental predictive power of aircraft counts (network size), total RC (number of edges) and the centralisation index (network structure) was examined in a sequential manner. First, all control variables were entered into step 1. Then, aircraft counts, RC and the centralisation index were entered one by one to test H1–H3 hierarchically (step 2–4). We also calculated the proportional reduction in variance (PRV) of each model, which is an effect size estimate that can be used in multilevel analyses (Raudenbush and Bryk 2002). Δ PRV was used to show the additive effect of each newly entered variable over and above previous variables. All predicting variables were group centred to show a constant intercept. The details of the analyses are shown in Table 2.

2.2.1. The effect of control variables

According to the coefficients estimated in step 1, perceived workload was higher when minimum separation was smaller ($\gamma_{10} = -.283, p < .001$) and there were more potential violations ($\gamma_{40} = .298, p < .05$). Participants completed the task slower when the average minimum separation was smaller ($\gamma_{10} = -3.187, p < .001$) and when there were more potential violations ($\gamma_{40} = 5.13, p < .01$). These findings are generally in accordance with previous research.

2.2.2. The test of hypotheses H1–H3

In a series of hierarchical models, we tested the incremental predictive power of the three testing variables (number of aircraft, total RC and centralisation) over and above previous steps. At step 2, more aircraft was found to result in higher perceived workload ($\gamma_{50} = 1.36, p < .001$) and longer task completion time ($\gamma_{50} = 18.86, p < .001$). H1 was fully confirmed. In the next step, higher total RC was found to result in higher perceived workload ($\gamma_{60} = .507, p < .001$) but its influence on task completion time was only approaching significance ($\gamma_{60} = 3.42, p = .10$). These findings partially confirmed H2. In the final step, it was found that centralised networks resulted in lower perceived workload ($\gamma_{70} = -3.38, p < .01$) as well as shorter task completion time ($\gamma_{70} = -47.11, p < .01$) suggesting that the more star-like a network is, the easier it is to handle. These findings fully confirmed our H3.

2.4. Discussion of Study 1

Study 1 used a static design directly following Boag et al. (2006). As our framework was built upon their work, using such a similar design can clearly show the incremental predictive power of the network characteristic. It provides initial evidence that network structure can influence operators' workload in addition to other important workload contributors (number of

Table 1. Zero-order correlations of all variables in both studies.

	Mean	SD	WD	TC	CD	VC	No	RC	Centralisation	PW1	PW2
WD	5.28	2.54	–								
TC	515.16	79.83	–.21**	–							
CD	43.92	32.63	–.31**	.29**	–						
VC	1.39	0.83	–.13**	.31**	.16**	–					
No	6.22	1.48	.37**	.45**	–.16**	.20**	–				
RC	6.28	3.48	–.35**	.51**	.20**	.14**	.62**	–			
Centralisation	.51	.22	–.10*	.20**	.39**	.12**	–.33**	–.36**	–		
PW1	16.46	4.98	–.14**	.03	.00	.06	.18**	.32**	–.27**	–	
TCT1	99.41	71.85	–.12**	.06	.04	.08*	.19**	.29**	–.21**	.35**	
PW2	17.48	3.57	–.12**	.01	.04	.08	.10*	.22**	–.18**	–	–
TCT2	83.26	43.78	–.03	.12**	–.01	.10*	.25**	.30**	–.24**	–	.33**

Note: WD, weighted distance of minimum separation; TC, time to conflict, CD, conflict duration; VC, violation counts; No., number of flights; RC, total relational complexity; PW1/2, perceived workload of Study 1 or Study 2; TCT1/2, task completion time of Study 1 or Study 2.

* $p < .05$; ** $p < .01$.

Table 2. HLM results predicting perceived workload and reaction time measures in Study 1.

Parameters	DV = Perceived Workload				DV = Task Completion Time			
	M1	M2	M3	M4	M1	M2	M3	M4
<i>Control variables</i>								
Intercept	16.46(.50)***	16.46(.50)***	16.46(.50)***	16.46(.50)***	99.41(6.72)***	99.41(6.72)***	99.41(6.72)***	99.41(6.72)***
Minimum separation	-.283(.063)***	-.628(.085)***	-.149(.113)	-.224(.110)*	-3.187(.917)***	-7.96(1.25)***	-4.78(2.69) ⁺	-5.82(2.92)*
Time to conflict	.000(.002)	-.014(.002)***	-.013(.002)***	-.009(.002)***	.024(.029)	-.170(.039)***	-.161(.040)***	-.109(.035)**
Conflict duration	-.008(.005) ⁺	.005(.004)	-.005(.005)	.004(.006)	-.038(.057)	.137(.062)*	.071(.074)	.197(.100)*
Violation counts	.298(.151)*	.007(.155)	.337(.154)*	.327(.153)*	5.13(1.98)**	1.10(2.05)	3.29(2.27)	3.17(2.28)
<i>Testing variables</i>								
Aircraft count		1.36(.14)***	.202(.231)	.289(.239)		18.86(2.69)***	11.15(6.04) ⁺	12.36(6.31)*
Relational complexity			.507(.104)***	.334(.126)*			3.42(2.13) ⁺	.90(2.70)
Network centralisation				-3.38(1.08)**				-47.11(15.50)**
PRV	.028	.162	.199	.216	.022	.138	.144	.158
△PRV	.028	.134	.036	.017	.022	.115	.006	.014

Note: $N = 18$ scenarios (Level 1) performed by 36 trainee controllers (Level 2); Parenthetical values indicate standard errors.

⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

flights and RC levels) identified in previous research. Furthermore, as early interventions are less likely to influence the performance of static tasks, the no-need-to-change-reference benefit in visual search, the first mechanism we have argued in the introduction, seems to be more feasible to explain the significant effect of network structure. However, it is necessary to see whether this benefit can still be observed in a more dynamic environment in which early intervention is more likely to happen. Therefore, we conducted Study 2 to further test these hypotheses.

3. Study 2

Study 2 was conducted to replicate and expand the findings of Study 1 by using a dynamic task design that is closer to real work environments. As controllers were able to make interventions at any time in performing the scenarios in this dynamic setting, they could take advantage of the centralised network structure in both the conflict detection phase (as Study 1 proved) and the conflict resolution process (by proactively resolving potential conflicts in advance). Therefore, the centralisation index can still be a significant predictor of controllers' workload.

3.1. Methods

3.1.1. Participants

Another 29 trainee controllers from the same training centre were recruited for this study by offering the same conditions as in Study 1. All participants were male between the ages of 21 and 24 ($M = 22.76$, $SD = .91$).

3.1.2. Tasks and procedure

The same research platform (interface and supportive tools) and scenarios used in Study 1 were used in this study. Participants were also asked to detect, mark and resolve the potential conflicts and when they believed that all potential conflicting pairs had been handled properly they could press the 'Finish' button. After they had resolved all potential conflicts, the same four questions as in Study 1 were asked again to measure perceived workload. However, there were several differences: (1) in each scenario, all aircraft appeared at the start, moving in a dynamic manner. The flight information (position, speed, altitude, etc.) was updated every second on the radar screen; (2) participants were able to resolve particular conflicts at any time including any potential future conflicts they spotted. Therefore, we anticipated that the participants would do the task faster and report fewer conflicts in each scenario. The summed scores of the four workload items were also used. We calculated the α coefficients of the workload scale for all 18 scenarios. The minimum value among the 18 values was .72, the maximum was .93 and the average was .85.

3.2 Results

On average, participants reported 1.76 conflicts ($SD = .59$) and spent 83.26 seconds completing each scenario ($SD = 43.8$). The final miss rate was 21.4%. The correlation between mean reaction time and miss rate ($r = .01$, *ns.*) suggests that there was no speed-accuracy trade-off.

We conducted similar multilevel analyses as in Study 1. Perceived workload and task completion time were treated as dependent variables. The ICCs of the null model suggested that 40.3% of the variance in perceived workload and 27.4% of the variance in task completion time were between individuals. This again justified our usage of multilevel analysis. The details of these further analyses are shown in [Table 3](#).

3.2.1 The effect of control variables

According to the coefficients estimated at step 1, perceived workload was higher when WD of minimum separation was smaller ($\gamma_{10} = -.170$, $p < .01$), and there was more potential violations ($\gamma_{40} = .337$, $p < .01$); task completion time was longer when TC was longer ($\gamma_{20} = .057$, $p < .01$), CD was shorter ($\gamma_{30} = -.071$, $p < .05$) and potential violations were more ($\gamma_{40} = 4.18$, $p < .01$). Except for TC, these findings are generally in accordance with previous research.

3.2.2 The test of hypotheses H1-H3

At step 2, more aircraft count was found to result in higher perceived workload ($\gamma_{50} = .649$, $p < .001$) and longer task completion time ($\gamma_{50} = 10.29$, $p < .001$). H1 was fully confirmed. In the next step, higher total RC was found to result in higher workload ($\gamma_{60} = .295$, $p < .01$) as well as longer task completion time ($\gamma_{60} = 4.63$, $p < .001$). These findings fully

Table 3. HLM results predicting perceived workload and reaction time measures in Study 2.

Parameters	DV = Perceived Workload				DV = Task Completion Time			
	M1	M2	M3	M4	M1	M2	M3	M4
Intercept	17.48(0.43)***	17.48(0.43)***	17.48(0.43)***	17.48(0.43)***	83.26(4.50)***	83.31(4.50)***	83.31(4.50)***	83.31(4.50)***
<i>Control variables</i>								
Minimum separation	-.170(.051)**	-.334(.060)**	-.059(.067)	-.097(.067)	-.258(.608)	-2.85(.73)**	1.46(.95)	.781(.898)
Time to conflict	-.002(.001)	-.009(.002)**	-.008(.002)***	-.006(.002)**	.057(.017)**	-.050(.023)*	-.036(.023)	-.003(.027)
Conflict duration	.00(.004)	.006(.004)	.000(.004)	.005(.004)	-.071(.033)*	.026(.036)	-.064(.038) +	.018(.039)
Violation counts	.337(.115)**	.199(.127)	.389(.115)**	.384(.115)**	4.18(1.55)**	1.99(1.61)	4.96(1.65)**	4.88(1.65)**
<i>Testing variables</i>								
Aircraft count		.649(.124)***	-.018(.132)	.027(.131)		10.29(1.80)***	-.160(2.16)	.63(2.07)
Relational complexity			.295(.085)**	.202(.083)*			4.63(.93)***	2.98(.91)**
Network centralisation				-.175(.60)**				-.30.86(9.83)**
PRV	.029	.094	.120	.129	.022	.113	.148	.165
△PRV	.029	.065	.026	.009	.022	.091	.035	.017

Note: $N = 18$ scenarios (Level 1) performed by 29 trainee controllers (Level 2); Parenthetical values indicate robust standard errors.
 + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

confirmed H2. In the final step, more centralised networks were found to result in lower level of perceived workload ($\gamma_{70} = -1.75, p < .01$) and faster task completion time ($\gamma_{70} = -30.86, p < .01$). These findings again confirmed H3.

3.3 Discussion of Study 2

In addition to confirming the predictive power of aircraft count (H1) and total RC (H2), Study 2 provides further evidence that the centralised networks are easier to handle and the network centralisation index can be used to predict operators' workload and task completion time. While the effect of most control variables was consistent with previous research, it is worth noting that longer time proximity was found to be related to longer task completion time. As our scenarios involved more aircraft scattered in different locations on the radar screen, one possible explanation is that: first, a longer time to reach the minimum distance is related to longer aircraft trajectories that need to be inspected; second, along these trajectories, other aircraft can appear (more likely when the aircraft are many) and this can interfere with the visual inspection by capturing and distracting operators' attention; third, this interference is stronger in Study 2 than in Study 1 because moving objects can distract people's attention more easily.

4. Discussion

Based on the RC metric developed by Boag et al. (2006), we created a RC network to analyse the effect of higher level interactional properties of aircraft on controllers' mental processes. While there have been some successful examples using network analysis in the domain of ergonomic research, it is quite rare to use such method to model the cognitive representation of the interaction of moving objects (aircraft). In this way, our incorporation of RC into a network is therefore a novel approach in this field.

Specifically, we discussed the influence of network centralisation on ATCos' mental workload associated with conflict detection and resolution. Across both a static and a dynamic experiment, we provided evidence suggesting that this network property can have a significant impact on controllers' mental processes in addition to aircraft count, RC and other control variables. Several interesting findings are worth discussing.

We first found that the summation of dichotomised RC was an effective predictor of controllers' workload and response time. Since Boag et al. (2006) only used three aircraft in their study, our research findings provide further validity for their metric as this variable added more predictive power over other control factors including air traffic count. It is also important to note that although we used a dichotomised coding of their original metric, their original ordinal metric can be readily incorporated into the network framework using a weighted network in future studies.

Second, we found consistent evidence that more centralised networks are easier and faster to handle when other important predictors were controlled for. This finding suggests that even when two sectors contain the same number of aircraft and total RC, they may still exert different cognitive demands on controllers. Therefore, using a linear aggregation of aircraft-level and even dyad level properties in order to get an overall workload indicator may not be enough and may even be misleading if we do not consider the overall network patterns.

We also proposed two possible mechanisms to explain why centralised networks may lessen the cognitive workload: (1) it reduces the need to update reference in conflict detection and (2) it provides a decision shortcut in conflict resolution. As controllers had to detect all potential conflicts in static tasks (Study 1) before making any interventions, and the total detection time was of major concern, the significant effect of centralisation found in this task suggests that it has a benefit in conflict detection. In the dynamic task (Study 2), participants performed faster and reported less-detected conflicts compared with the static task. We believed that this is because they can proactively resolve potential conflicts in advance in performing the task. Therefore, the significant effect of the centralisation index suggests that intervention decision-making can also take benefit from the centralised network structure. Future studies may benefit from measuring both processes separately in a more precise manner.

We can now suggest several possible theoretical and practical contributions of our research as well as future research directions. First, our framework has great potential in the currently underdeveloped modelling of the dynamic decision-making processes of ATCos (Loft et al. 2007; Neal et al. 2014). For example, it is possible that controllers' initial attention is allocated in a random or a habitual manner (Fothergill and Neal 2013; Neal and Kwantes 2009). But after scrutinising the relations among aircraft, controllers may find that using the aircraft with high centrality (and its trajectory) as a template may enhance future conflict detection and intervention. If this is the case, the network property can be used as an attention organisation principle in more fundamental cognitive processes and be incorporated within formal computational models of cognition (e.g. Fu et al. 2014).

In addition, this framework can be integrated with the dynamic system approach. First, the dynamic system approach may benefit from changing its 'god-like-view' of purely physical relationships to its impact on cognition (e.g. RC). Second,

the RCN framework can be expanded by a multiplex network approach to understand how different relational attributes interact to influence controllers' decision-making. In a multiplex network, the relationship between two objects is a vector rather than a single value. For example, aircraft with small lateral distance (close) can be perceptually grouped together which can be processed as a chunk to save attention resources (Landry, Sheridan, and Yufik 2001), however, this can only be possible when there were no potential conflicts among them ($RC = 0$ condition). Therefore, the match between the two networks (RC network and a closeness network) can serve as an indicator of whether there could be some chance to reduce cognitive burden. Third, some dynamic network properties such as network entropy can be used to expand the current RCN framework. For example, network entropy evaluates the randomness and predictability in a dynamic network without losing its structural information. We could imagine that a network with a continuous changing and unpredictable structure can be more demanding than a less variable one.

Moreover, it is also possible that controllers may have a particular way to process some specific network structures (e.g. star-like ones). As humans do have the ability to directly process global patterns in a faster heuristic way that is less bounded by limited working memory capacity, it is possible that controllers may process star-like networks using pattern recognition rather than a step-by-step analysis. If this is the case, we are more likely to observe the difference between a star-like and a chain-like network in more experienced controllers because they are more adept in processing the global pattern from the very beginning by finding the most pivotal node in a fast and accurate manner. Future research may test these possibilities on controllers with different levels of experience. In addition, from a cultural difference perspective, the cultural background of the participants in our research (Chinese) is generally considered to promote a higher inclination to use holistic rather than analytical processing compared to the western culture (Nisbett et al. 2001). It might therefore be interesting to investigate whether this finding can be replicated among controllers from the western countries, and thus explore a phenomenon that is not often addressed in human factor research.

While our research focused on the global network properties (centralisation), future research may also benefit from examining the network properties of each node (the position in the network). It is important to note that we only investigated the effect of network centralisation at the starting point in each scenario. In a real-time task environment such as Study 2, aircraft movement and controllers' early interventions may change the value of such a metric in a dynamic manner. As a result, the predictive power of the proposed metric may actually have been underestimated in our research. This is because the centralisation metric can actually evaluate the dynamic fragility or vulnerability of a network. For example, it has been found that many real-world networks are very vulnerable to targeted deletion of nodes with high centrality (Albert, Jeong, and Barabási 2000). In the RCN framework, it is therefore important to know the actual intervention sequence adopted by controllers. Some initial evidence has been found that aircraft at the centre of the RC network (with higher centrality) is given higher priority in conflict resolution (Zhang, Ren, and Wu 2014). Further, aircraft's network positions may interact with other complexity factors to influence the cognitive processes of ATCos. For example, the difficulty caused by the controllability of an aircraft (when an aircraft is short of fuel or in other kinds of emergency states, its flight parameters are less controllable by ATCos) may depend on the aircraft's position in the RC network. If an aircraft is at the network centre, its uncontrollability would cause a high workload as ATCos must monitor all other related aircraft to resolve potential conflicts. On the other hand, when the aircraft is at the margin of the network, difficulty would not increase substantially even if facing controllability issues.

Finally, the network framework may also be used to develop intelligent supporting tools to improve human-computer interaction in conflict detection and resolution. It has been suggested that providing relational information directly to the controllers can be very helpful (Vuckovic et al. 2013), but it may in fact be more helpful to add some network-based information. As our findings showed, it is time-consuming for controllers to handle chain-like patterns. It is therefore possible that if certain help were given (e.g. offering possible solutions) in these situations, controllers' response times might be greatly reduced. Future research is needed to see whether this approach might be naturally and consistently integrated with controllers' current work practices.

Several limitations must be addressed before concluding. First, while the proposed network approach provides an effective way to quantify complex aircraft interactional patterns, scenarios as complex as in our experiments might not be common in real-life circumstances. This is because most current ATCos usually monitor aircraft along a few predefined routes which restrict potential RC – only in some very busy areas or abnormal situations (bad weather, military aircraft, etc.) are complex interactions likely to occur. However, the implementation of closer separation standards and free flights in future ATM systems such as SESAR and NextGEN may require controllers to deal with more aircraft interacting with a higher degree of freedom. In such circumstances, the network approach may be more useful.

It also needs to be noted that the effect sizes we found in our study were significant but quite small (total effect: 13–22%; incremental effect of RC: .6–3.6%, and centralisation: .9–1.7%). In the study by Boag et al. (2006), the independent variables explained 30–50% of total variance. However, in their research, the dependent variables were averaged across different participants to create one value for each scenario (a total of 60 scenarios). This averaging method

certainly found a very high effect size because the between-individual errors had already been removed. Indeed, using such procedure to analyse our data, we obtained a much higher effect size (59–80%) and a significant incremental effect size of RC (13–18%) as well as network centralisation (6–9%). However, as the sample size equals to the number of scenarios in using such method and we used relatively few scenarios (18 in total) in our research, the effect size might be inflated. Future research may benefit from using more scenarios to get a clearer picture of the effect sizes.

Finally, it is worth noting that, our research only investigated conflict detection and resolution. While these tasks are very important for ATCOs, they also need to accept and handoff aircraft, provide weather advice to pilots and coordinate with other professionals to regulate the airflow. Therefore, future studies may take these types of workload into consideration.

5. Conclusion

Although it is not new to use network analysis in the research of basic psychology (Baronchelli et al. 2013), as well as ergonomics (Stanton 2014; Walker et al. 2010), to the best of our knowledge, our study is the first to implement such an approach in ATC conflict detection and resolution tasks. While the generalisability of this framework requires future exploration, our research provides initial evidence that it can be used in an ATC context. The relational complexity network framework provides a new tool to link the holistic pattern of aircraft interactions with cognitive processes and computational representations.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the National Basic Research Program of China [grant number: 2011CB302201], the Key Research Program of the Chinese Academy of Sciences [grant number: KJZD-EW-L04] and the Natural Scientific Foundation of China [grant number: 31300875 and 31070915].

Notes

1. Email: jjzhongyang@msn.com.
2. We used the binary coding instead of the 0–4 coding system used by of Boag et al. (2006) for the following reasons. First, we believe that it is mathematically feasible to use the ordinal metric and a weighted network approach. However, as more different patterns (such as degree distribution) may occur in a weighted network, it requires more varied scenarios to offer empirical evidence which can be very difficult for researchers to design and also burdensome for participants. It might be more practical to use such a framework if certain ‘big data’ from the real ATC operation centres can be obtained. Second, most of the $RC > 0$ conditions in our scenarios (94%) were actually the $RC = 2$ conditions according to the original Boag et al. (2006) metric. In this regard, we believe that this dichotomous coding would not reduce much information.
3. These filler scenarios were used for testing a different hypothesis; however, including them in the current analysis does not change the major findings. Interested readers may contact the authors for further information about the analysis including these scenarios.
4. This is the ratio between pairs still in conflict after they provided resolution and pairs in conflict at the start of scenarios. The conflict detection rate (pairs reported to be in conflict using the detection function), on the other hand, was 40.4% for Study 1 and 49.8% for Study 2. They were much higher than the three-aircraft conditions in Boag et al. (2006) (about 10%) using experienced controllers but similar to the findings in an earlier study (about 40.0%) (Landry, Sheridan, and Yufik 2001) using amateurs and more aircraft. It must be mentioned that our tasks were relatively more difficult since: (1) nearly all conflicting pairs involved both horizontal conflict and altitude traversing; (2) the conflict occurrence rate was quite low (8%), and (3) the participants were not able to see their detection results, therefore they may have forgotten the data that had already been entered into the system. More importantly, in many of our scenarios, one aircraft can cause many potential conflicts, reporting one of them can be considered by participants to be enough as resolving it can prevent all other related conflicts. This is especially the case for Study 2 in which early intervention can prevent some pre-arranged conflicts to appear at all. If we adopt such a lenient criterion by treating detecting one pair among all related pairs (sharing one airplane in common) as a hit for them all, then the detection miss rate will sharply drop to 22.6% and 38.3% for Studies 1 and 2, respectively. If we further consider resolution, the rate will drop to 14.6% and 21.4%, respectively. None of them is correlated with completion time.

References

- Albert, R., H. Jeong, and A.-L. Barabási. 2000. “Error and Attack Tolerance of Complex Networks.” *Nature* 406 (6794): 378–382.
- Anderson, P. W. 1972. “More is Different.” *Science* 177 (4047): 393–396.
- Athènes, S., P. Averty, S. Puechmorel, D. Delahaye, and C. Collet. 2002. “ATC Complexity and Controller Workload: Trying to Bridge the Gap.” In *Proceedings of the International Conference on HCI in Aeronautics*. 56–60. <http://www.aaai.org/Papers/HCI/2002/HCI02-009.pdf>

- Baronchelli, A., R. Ferrer-i-Cancho, R. Pastor-Satorras, N. Chater, and M. H. Christiansen. 2013. "Networks in Cognitive Science." *Trends in Cognitive Sciences* 17 (7): 348–360. doi:10.1016/j.tics.2013.04.010.
- Boag, C., A. Neal, S. Loft, and G. S. Halford. 2006. "An Analysis of Relational Complexity in an Air Traffic Control Conflict Detection Task." *Ergonomics* 49 (14): 1508–1526.
- Borgatti, S. P., M. G. Everett, and L. C. Freeman. 2002. "Ucinet for Windows: Software for Social Network Analysis." <http://www.citeulike.org/group/11708/article/6031268>
- Delahaye, D., and S. Puechmorel. 2000. "Air Traffic Complexity: Towards Intrinsic Metrics." Presented at the 3rd USA/Europe Air Traffic Management R&D Seminar, Napoli, Italy.
- Delahaye, D., and S. Puechmorel. 2010. "Air Traffic Complexity Based on Dynamical Systems." Presented at the 49th IEEE Conference on Decision and Control, IEEE, Atlanta, USA. 2069–2074. doi:10.1109/CDC.2010.5718004.
- Fothergill, S., and A. Neal. 2013. "Conflict-Resolution Heuristics for En Route Air Traffic Management." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 57: 71–75.
- Freeman, L. C. 1978. "Centrality in Social Networks Conceptual Clarification." *Social Networks* 1 (3): 215–239.
- Fu, X., L. Cai, Y. Liu, J. Jia, W. Chen, Y. Zhang, G. Zhao, Y. J. Liu, and C. Wu. 2014. "A Computational Cognition Model of Perception, Memory, and Judgment." *Science in China Series F: Information Sciences* 57 (3): 032114:1–032114:15.
- Gianazza, D. 2010. "Forecasting Workload and Airspace Configuration with Neural Networks and Tree Search Methods." *Artificial Intelligence* 174 (7–8): 530–549. doi:10.1016/j.artint.2010.03.001.
- Gobet, F. 1997. "A Pattern-Recognition Theory of Search in Expert Problem Solving." *Thinking & Reasoning* 3 (4): 291–313.
- Hilburn, B. 2004. *Cognitive Complexity in air Traffic Control: A Literature Review*. Technical Report No. 04/04. Belgium: Eurocontrol.
- ICAO. 2005. *Global Air Traffic Management Operational Concept*. No. Doc 9854. Montreal, Canada.
- ICAO. 2013. *Global Air Transport Outlook to 2030 and Trends to 2040*. No. Cir 333, AT/190. Montréal, Canada.
- Kahneman, D., and G. Klein. 2009. "Conditions for Intuitive Expertise: A Failure to Disagree." *American Psychologist* 64: 515–526.
- Landry, S. J., T. B. Sheridan, and Y. Yufik. 2001. "A Methodology for Studying Cognitive Groupings in a Target Tracking Task." *IEEE Transactions on Intelligent Transportation Systems* 2 (2): 92–100.
- Laudeman, I. V., S. G. Shelden, R. Branstrom, and C. L. Brasil. 1998. *Dynamic Density: An Air Traffic Management Metric*. No. NASA-TM-1988-11226. Moffet Field, CA: NASA Ames Research Center.
- Loft, S., S. Bolland, M. S. Humphreys, and A. Neal. 2009. "A Theory and Model of Conflict Detection in Air Traffic Control: Incorporating Environmental Constraints." *Journal of Experimental Psychology: Applied* 15 (2): 106–124.
- Loft, S., P. Sanderson, A. Neal, and M. Mooij. 2007. "Modeling and Predicting Mental Workload in en Route Air Traffic Control: Critical Review and Broader Implications." *Human Factors* 49 (3): 376–399.
- Mogford, R. H., J. A. Guttman, S. L. Morrow, and P. Kopardekar. 1995. *The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature*. No. DOT/FAA/CT-TN95/22. Atlantic City, NJ: Federal Aviation Administration, William Hughes Technical Center.
- Neal, A., S. Hannah, P. Sanderson, S. Bolland, M. Mooij, and S. Murphy. 2014. "Development and Validation of a Multilevel Model for Predicting Workload Under Routine and Nonroutine Conditions in an Air Traffic Management Center." *Human Factors* 56: 287–305.
- Neal, A., and P. J. Kwantes. 2009. "An Evidence Accumulation Model for Conflict Detection Performance in a Simulated Air Traffic Control Task." *Human Factors* 51 (2): 164–180. doi:10.1177/0018720809335071.
- Nisbett, R. E., K. Peng, I. Choi, and A. Norenzayan. 2001. "Culture and Systems of Thought: Holistic Versus Analytic Cognition." *Psychological Review* 108 (2): 291.
- Olivers, C. N. L., J. Peters, R. Houtkamp, and P. R. Roelfsema. 2011. "Different States in Visual Working Memory: When it Guides Attention and When it Does Not." *Trends in Cognitive Sciences* 15 (7): 327–334.
- Prandini, M., L. Piroddi, S. Puechmorel, and S. L. Brazdilova. 2011. "Toward Air Traffic Complexity Assessment in New Generation Air Traffic Management Systems." *Intelligent Transportation Systems, IEEE Transactions on* 12 (3): 809–818.
- Rantanen, E. M., and A. Nunes. 2005. "Hierarchical Conflict Detection in Air Traffic Control." *The International Journal of Aviation Psychology* 15 (4): 339–362.
- Raudenbush, S. W., and A. S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed Thousand Oaks, California: Sage.
- Remington, R. W., J. C. Johnston, E. Ruthruff, M. Gold, and M. Romera. 2000. "Visual Search in Complex Displays: Factors Affecting Conflict Detection by Air Traffic Controllers." *Human Factors* 42 (3): 349–366.
- Simon, H. A. 1991. "The Architecture of Complexity." In *Facets of Systems Science*, edited by G. J. Klir, 457–476. New York: Springer.
- Stankovic, S., É. Raufaste, and P. Averty. 2008. "Determinants of Conflict Detection: A Model of Risk Judgments in Air Traffic Control." *Human Factors* 50 (1): 121–134. doi:10.1518/001872008X250584.
- Stanton, N. A. 2014. "Representing Distributed Cognition in Complex Systems: How a Submarine Returns to Periscope Depth." *Ergonomics* 57 (3): 403–418. doi:10.1080/00140139.2013.772244.
- Vuckovic, A., P. Kwantes, and A. Neal. 2011. "A Dynamic Model of Decision Making in ATC Adaptation of Criterion Across Angle and Time." . Presented at the Human Factors and Ergonomics Society 55th Annual Meeting, Las Vegas, Nevada. pp. 330–334.
- Vuckovic, A., P. Sanderson, A. Neal, S. Gaukrodger, and B. L. W. Wong. 2013. "Relative Position Vectors An Alternative Approach to Conflict Detection in Air Traffic Control." *Human Factors* 55: 946–964.
- Walker, G. H., N. A. Stanton, P. M. Salmon, D. P. Jenkins, and L. Rafferty. 2010. "Translating Concepts of Complexity to the Field of Ergonomics." *Ergonomics* 53 (10): 1175–1186. doi:10.1080/00140139.2010.513453.
- Zhang, J., J. Ren, and C. Wu. 2014. "Modeling Air Traffic Controllers' Decision Making Processes with Relational Complexity Network." Presented at the 17th International IEEE Conference on Intelligent Transportation Systems, Qingdao, China.

Appendix

Table A1. Basic configurations of the 18 scenarios used in the present study.

Scenario No.	Average WD (separation units)	Average TC (s)	Average CD (s)	VC	No. of flights	Relational complexity	Centralisation (%)
1	3.46	431.13	23.30	1	4	2	66.67
2	2.39	483.28	100.80	1	4	3	100.00
3	2.00	500.20	55.40	2	4	4	66.67
4	2.70	419.87	16.90	1	4	5	33.33
5	8.05	527.49	30.30	1	6	3	60.00
6	8.48	484.98	21.75	2	6	3	60.00
7	7.13	503.11	0.00	0	6	4	20.00
8	7.39	448.81	0.00	0	6	5	40.00
9	7.52	567.01	117.50	1	6	5	80.00
10	4.24	538.77	38.60	3	6	5	70.00
11	3.78	524.28	80.50	2	6	7	20.00
12	4.33	507.37	92.40	1	6	12	30.00
13	11.16	396.86	21.75	2	8	3	42.86
14	5.55	603.77	34.80	1	8	7	42.86
15	7.30	558.87	29.25	2	8	9	52.38
16	3.53	416.53	32.60	1	8	12	19.05
17	3.23	672.10	45.30	1	8	12	76.19
18	2.79	688.44	49.40	3	8	12	38.10

Note: WD, weighted distance of minimum separation; TC, time to conflict; CD, conflict duration; VC, violation counts; No., number of flights; RC, total relational complexity.