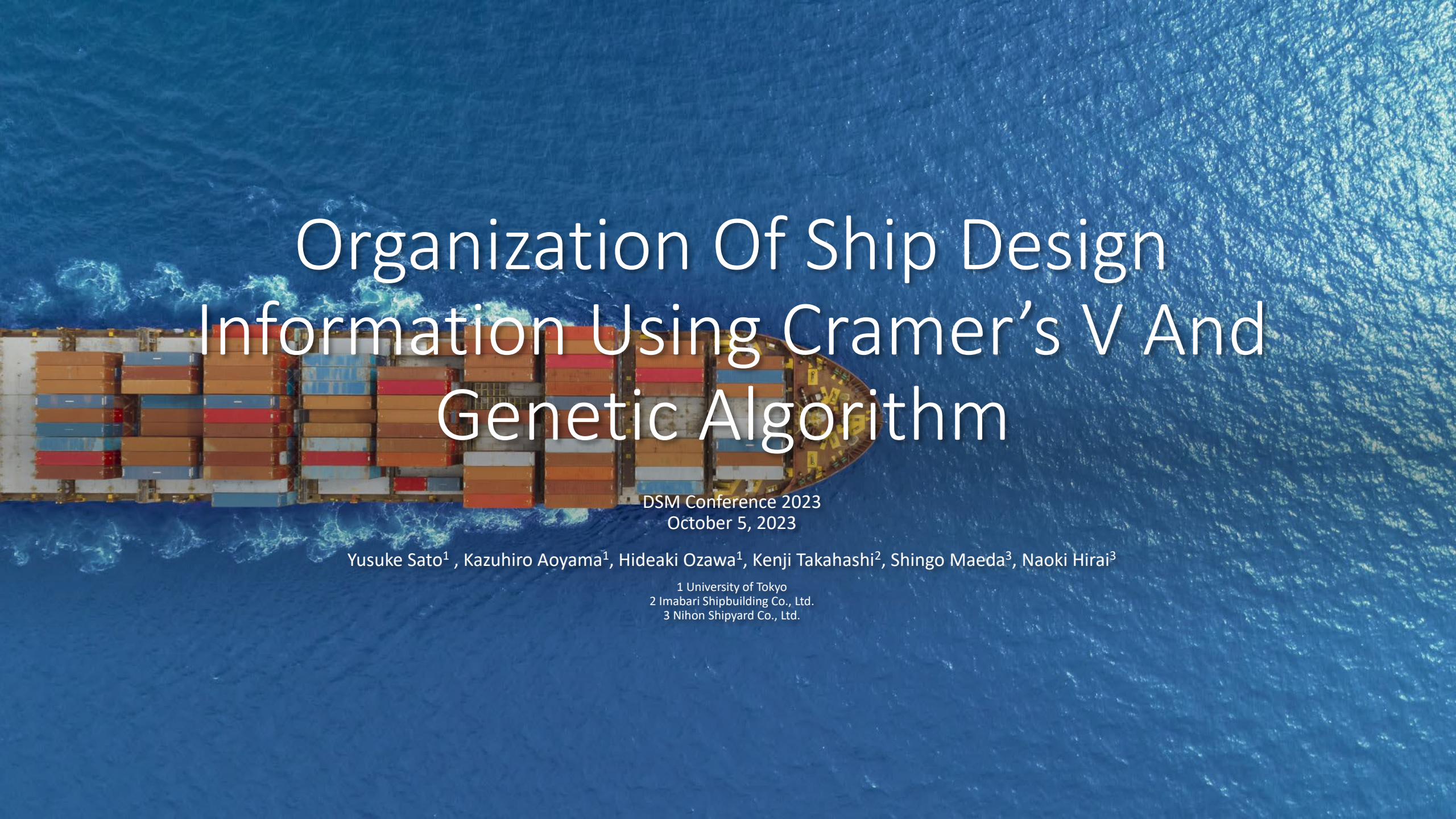


Organization Of Ship Design Information Using Cramer's V And Genetic Algorithm

DSM Conference 2023
October 5, 2023

Yusuke Sato[1] , Kazuhiro Aoyama[1], Hideaki Ozawa[1], Kenji Takahashi[2], Shingo Maeda[3], Naoki Hirai[3]

1 University of Tokyo
2 Imabari Shipbuilding Co., Ltd.
3 Nihon Shipyard Co., Ltd.

Contents

- Introduction
- Methods
- Results of applying the method to ships
- Conclusion

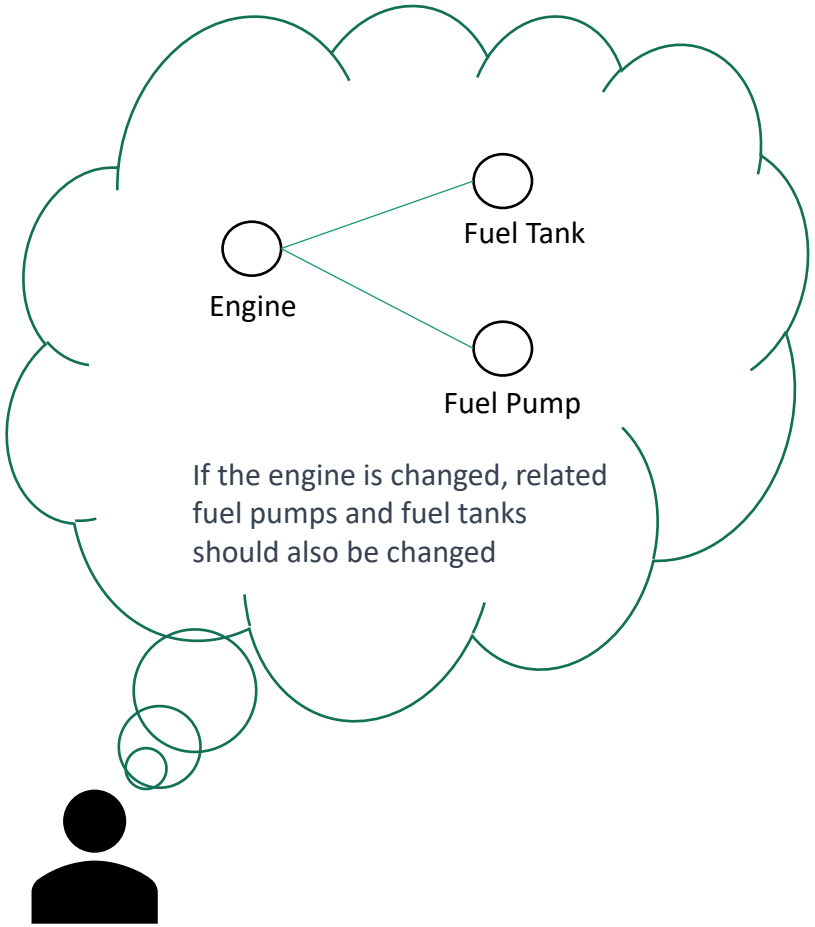
Contents

- Introduction
- Methods
- Results of applying the method to ships
- Conclusion

Backgrounds & Purpose

- Shipbuilding is custom-made and is designed by repeatedly changing components based on existing specifications.
- Specification changes are not limited to modifying a single component but often require adjustments to related surrounding components.
- It is important to understand related components for specification changes.
- These knowledge is not organized and specification changes rely on the intuition and experience of skilled workers.
- Coordination among departments in specification changes is also done on an ad hoc basis.

Extract information such as
"which specifications need to be changed at the same time" and
"whether the specification change is cross-departmental”
so that even non-skilled designers can design efficiently.



Contents

- Introduction
- Methods
- Results of applying the method to ships
- Conclusion

Outline of the Method

The method of this research can be divided into three steps:

Identification of components that are changed simultaneously

- Identify specifications that are changed simultaneously using Cramer's V, a measure of independence.

Check by designer

- Extracting relevant components from the designer's perspective from among components that have high Cramer’s V.
- Attempt to reduce the burden on designers

Analysis using DSM

- Describe the information among the components in the DSM and perform clustering.
- Analyze the impact of specification changes, etc.

Calculation of Cramer's V

- Ships are complex systems with many components, so it is difficult to manually enter all the relationship of components.
- Compare specification documents to find components that are changed simultaneously

	Ship A	Ship B	Ship C	Ship D	Ship E
Engine	x	x	x	y	y
Fuel Tank	a	a	a	b	b
Stern Tube	z	z	z	z	z

...

→ Shouldn't the engine and fuel tank be changed at the same time?

⋮

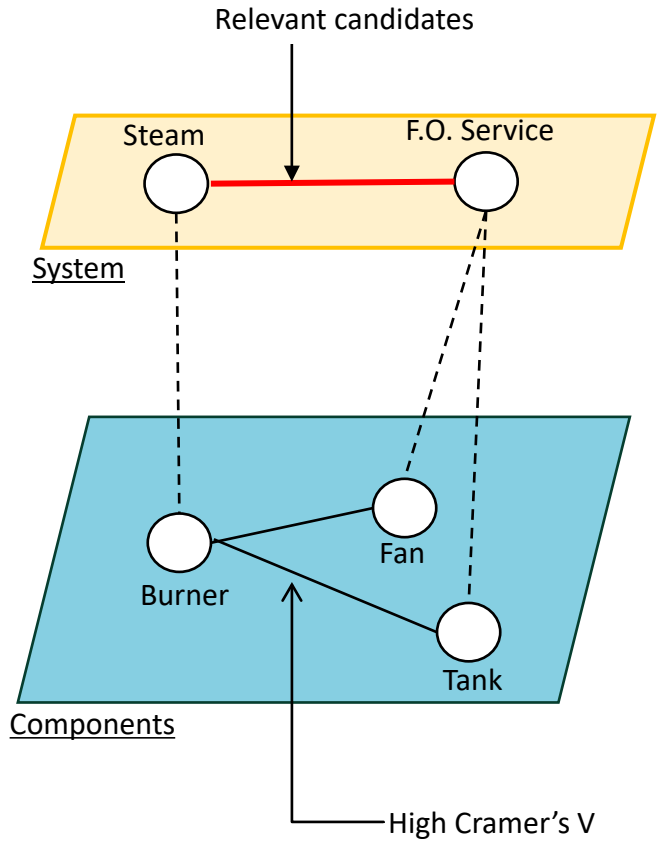
- Calculate the χ^2 test-based Cramer's V: $v = \sqrt{\frac{\chi^2}{n(k-1)}}$ for each component pair to determine if they are changed at the same time
  - n: number of samples, k: smaller number of kinds of two elements
  - Takes a value between 0 and 1
  - The closer to 1, the stronger the relationship between the two elements.
- Pairs with Cramer's V exceeding a defined threshold are considered components that are changed simultaneously.

Cramér, 1946. Mathematical Methods of Statistics. Princeton: Princeton University Press, 282.

Improved Accuracy of Information through Designer's Checks

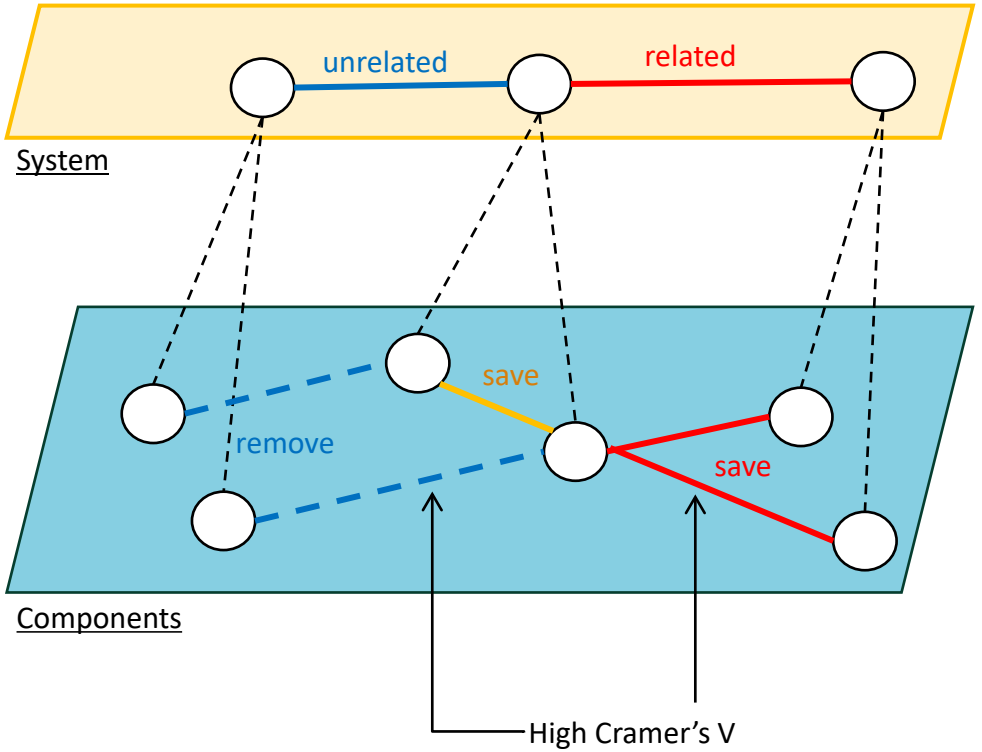
- Changed at the same time ≠ related for design
- Designer should correct the information.
- Checking all pairs of components is impractical.

- To perform a rough check first, we utilized a hierarchical structure of specification information.
  - several components belong to a system.
    (e.g. component "Fan" belongs to system "F.O. Service")
- If there is a high Cramer's V between components belonging to two different systems, those systems are considered as relevant candidates.



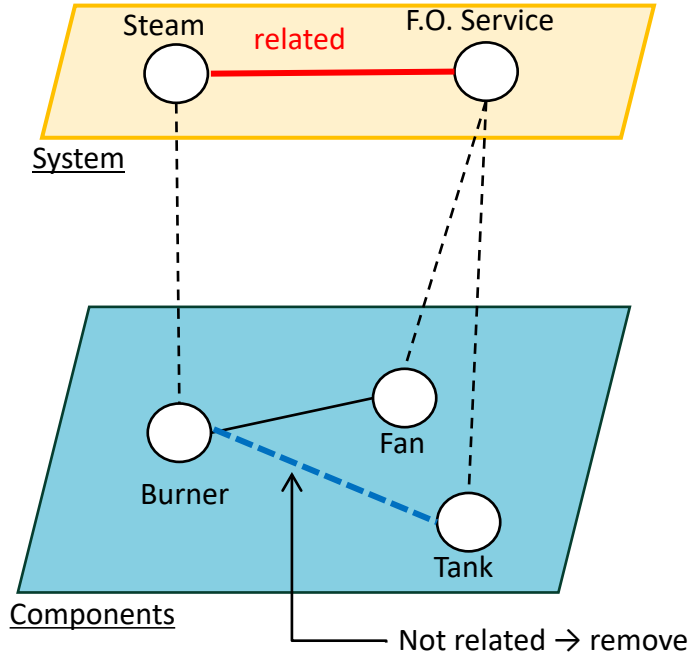
System-level Designer's Check

- Designer check information of the systems that are considered as relevant candidates to classify weather they are truly related or not.
- Reflect the result of check for components-level relationship
  → Classify the relationships of components into three categories:
  - Belonging to the same system
  - Belonging to related different systems
  - Belonging to unrelated different systems
- Remove the relationship between components belonging to unrelated systems
  → Remove information that becomes noise for the designer



Component-level Designer's Check

- The classification based on the systems is rough and contains inaccuracies.
- Check the relationships at the component level.
- Remove relationships between components that are not related for design.



DSM Clustering

- Describe information between components in DSM.
- Clustering use the IGTA+ algorithm, which is a genetic algorithm.
- Pairs of components belonging to different systems are given more weight than pairs belonging to the same system.

➢ If systems are considered as departments, this weighting encourages specification changes that require collaboration to be clustered together.

Borjesson, F, & Hölttä-Otto, K. "Improved Clustering Algorithm for Design Structure Matrix." *Proceedings of the ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 3, 38th Design Automation Conference, Parts A and B. 2012, pp. 921-930

Contents

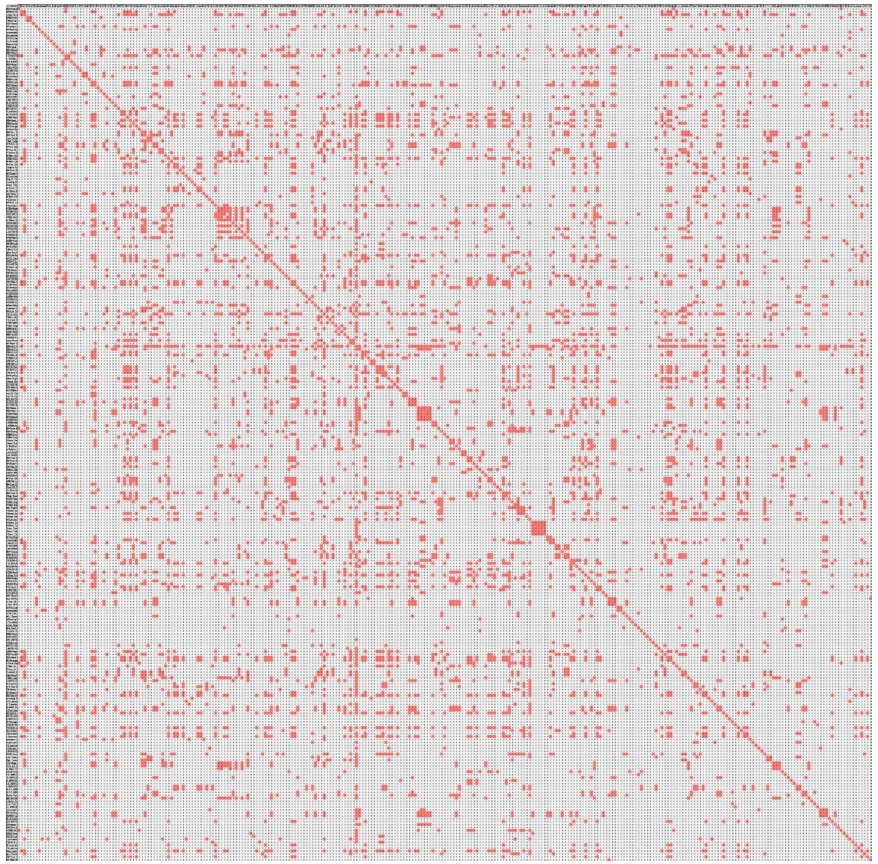
- Introduction
- Methods
- Results of applying the method to ships
- Conclusion

The Subject of Analysis for This Research

- Specification documents for 30 vessels of the 63,000/64,000 Mt BULK CARRIER series.
- There are 291 components related to machinery, electrical, and ship equipment.
  (excluding 39 components that exist only once out of the 330 components)
  - e.g. Main Engine, Auxiliary Air Compressor, Bilge & Ballast Pump.
- There are 30 systems, and each system includes several components.
  - e.g. Main Propulsion Unit, Compressed Air, Fire & Ballast.

Calculation of Cramer's V

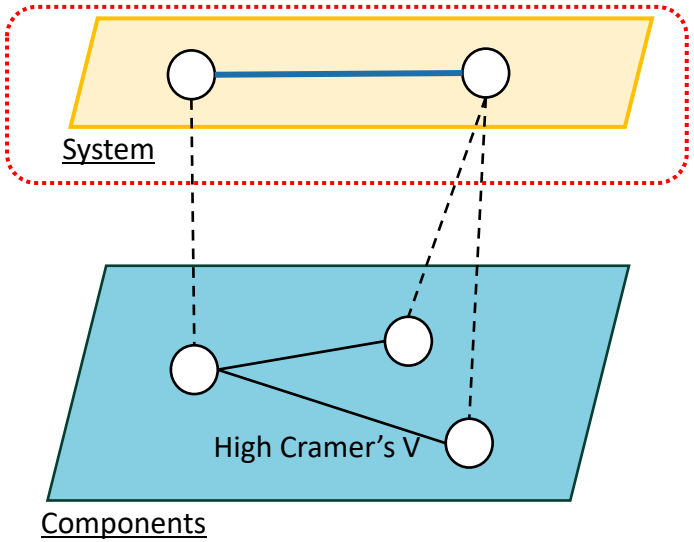
Calculate Cramer's V for all pairs of components and identify pairs of components that are changed simultaneously.



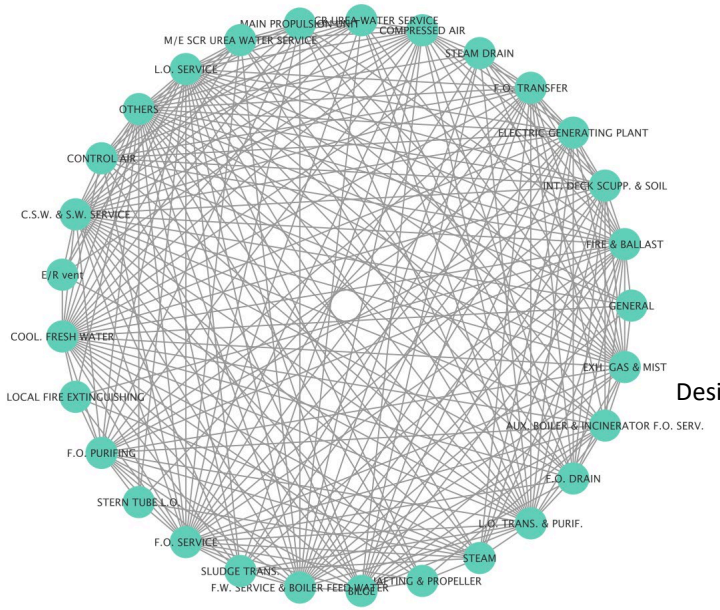
- Pairs of Components with value of Cramer's V of 0.96 or higher are defined as components that change simultaneously.
- Cells colored in the left matrix represent components that change simultaneously.
- Among the 42,195 pairs, 3,656 pairs of components, equivalent to 8.67%, met this criterion.

System-level Designer’s Check

From the information on components changed simultaneously, create information on system-level relationships.

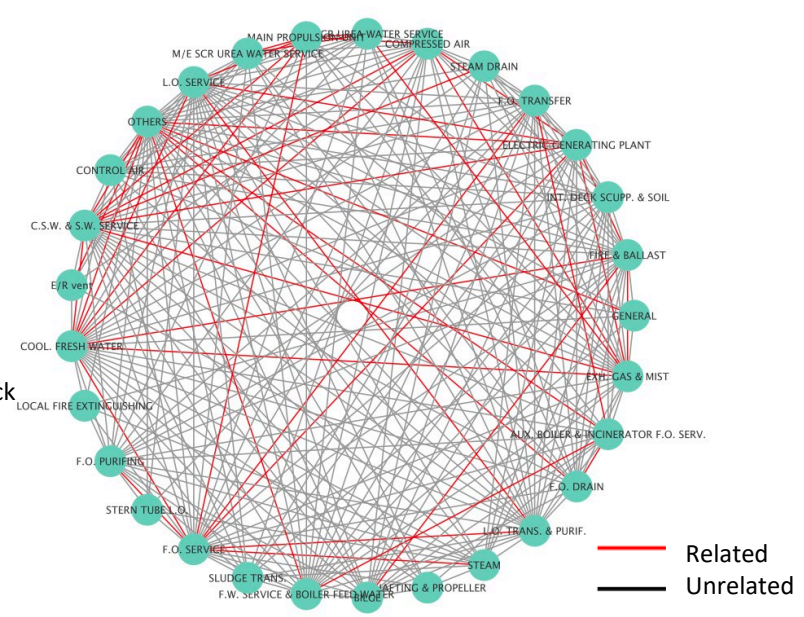


If there are pairs of components belonging to two different systems with a Cramer’s V of 0.96 or higher, consider the two systems as potential candidates for having a relationship.



Network representing candidates for related systems

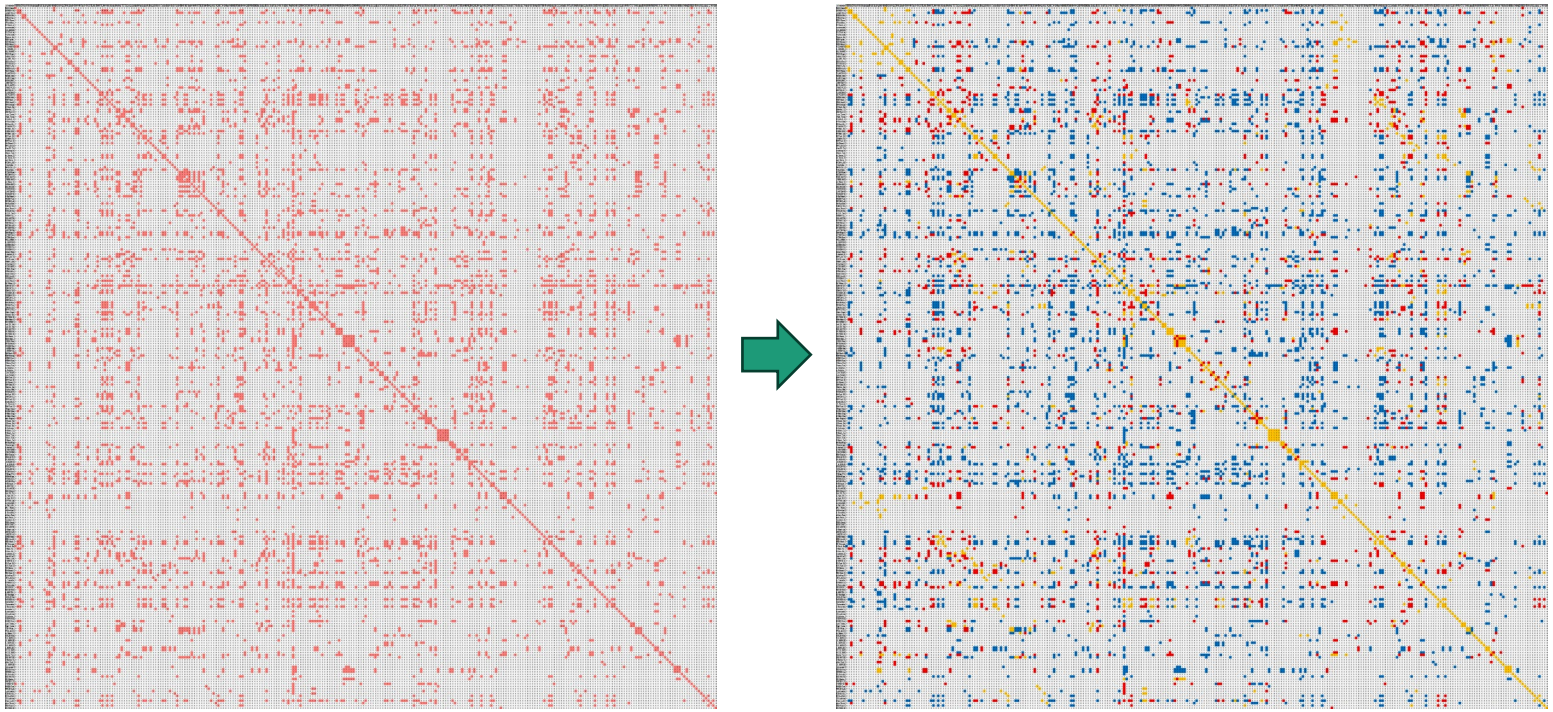
Designer's Check



Network representing systems that are related from a designer's perspective

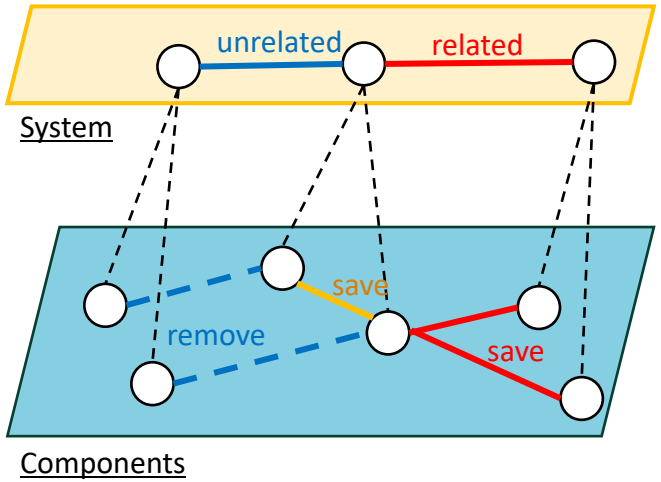
Reflect of System-level Check

Reflecting the system relationships onto the component relationships and remove relationships between components belonging to unrelated system.



Matrix indicating components with high Cramer's V

Matrix considering system relationships

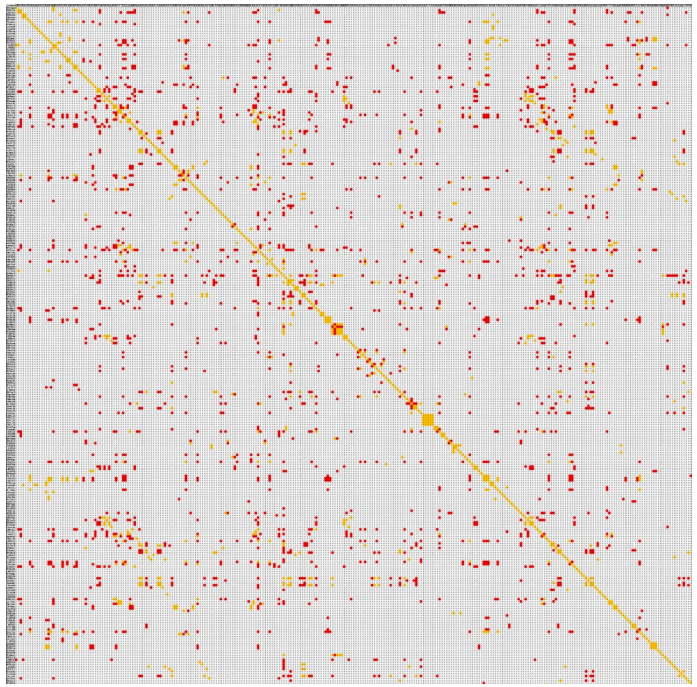


Referencing system-level relationship information, exclude pairs of components that change simultaneously but are not related for design.

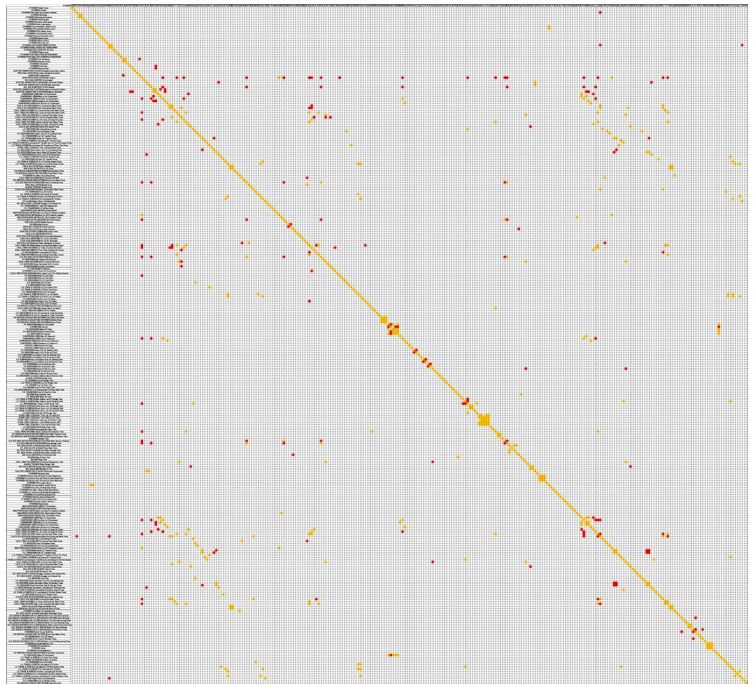
Belonging to the same system
Belonging to related different systems
Belonging to unrelated different systems

Component-level Designer's Check

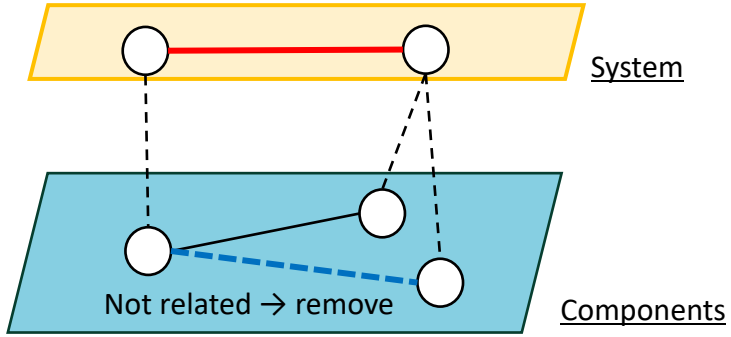
Obtain accurate information about components that are related for design and change simultaneously.



Matrix with relationships removed between components belonging to unrelated systems



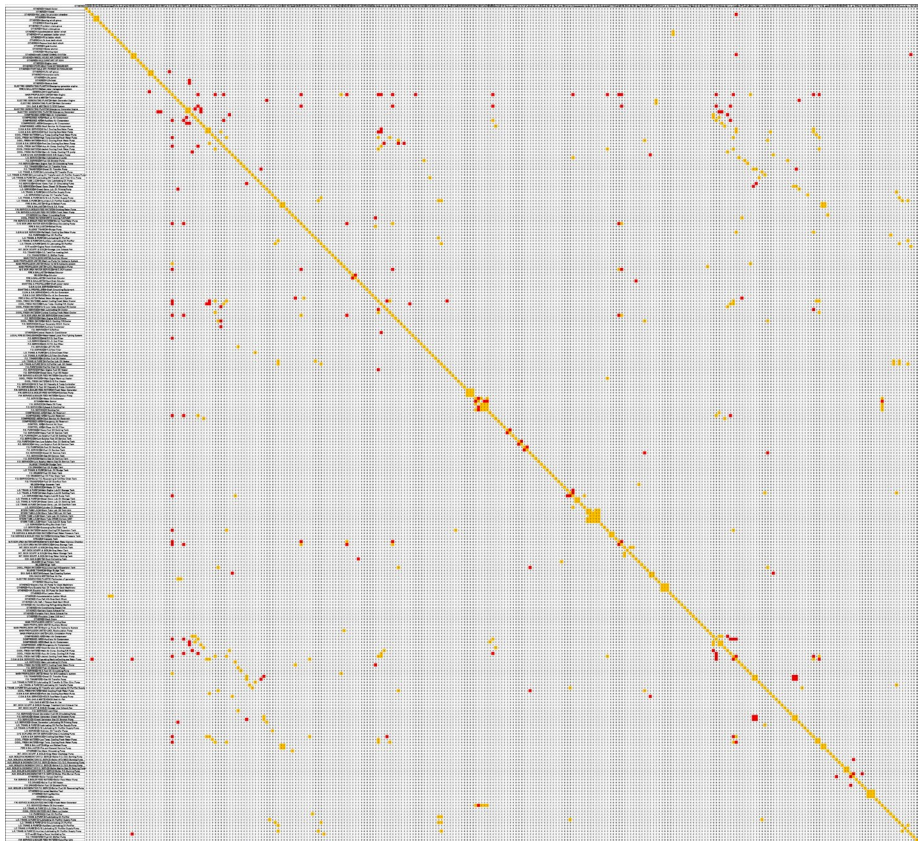
Matrix indicating the relationships between components that are changed simultaneously as obtained through designer's checks



- Remove pairs that had relationships at the system level but not at the component level.
- By conducting macro-level checks beforehand, reduce the amount of verification required at the component level.

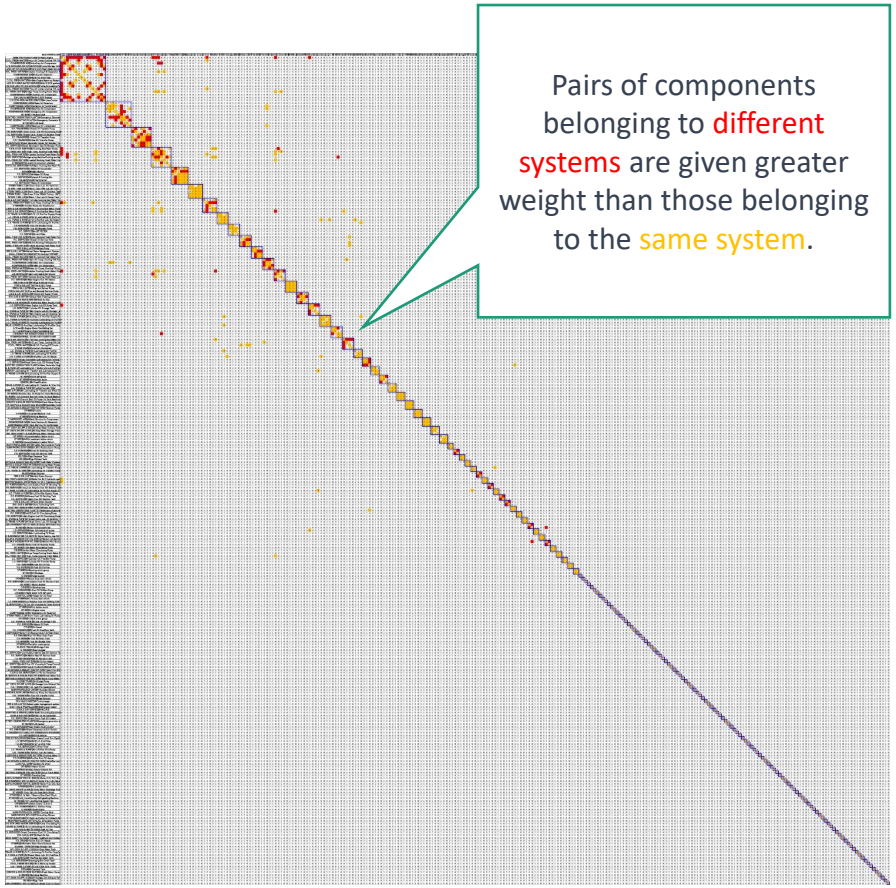
Belonging to the same system
Belonging to related different systems

DSM Clustering



Matrix identifying components with related relationships after designer's checks

Clustering by IGTA+



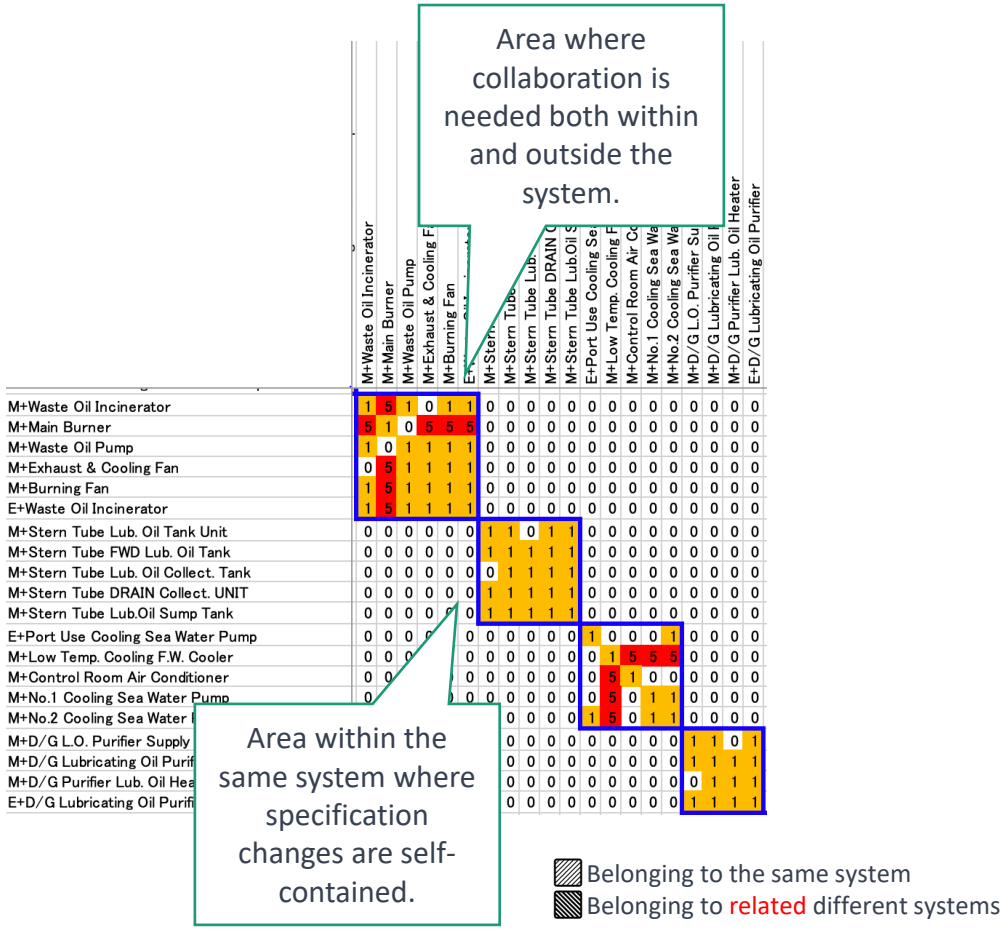
Matrix where related components have been grouped through clustering

Belonging to the same system
Belonging to related different systems

Borjesson, F, & Hölttä-Otto, K. "Improved Clustering Algorithm for Design Structure Matrix." *Proceedings of the ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 3, 38th Design Automation Conference, Parts A and B. 2012, pp. 921-930

Results of DSM Clustering

- Weight is assigned to relationships between components belonging to different systems rather than those within the same system.
  → It is possible to proactively identify specification changes that require collaborations.
- It is also possible to find components that can be designed independently of others.
- By reviewing each cluster (areas with strong relationships), it is possible to reduce rework during evaluation of specification changes.



Contents

- Introduction
- Methods
- Results of applying the method to ships
- Conclusion

Conclusion

Conclusion

- Through the use of Cramer's V and macro-level analysis, we were able to accurately identify components that should be changed simultaneously while reducing the effort of designers.
- DSM clustering allowed us to discover related components groups, potentially assisting in specification changes.

Remaining Challenges and Future Works

- Need for further efficiency improvements in the designer's check process.
- Cramer's V cannot handle causality relationship between components.

- Integration with different types of data, such as text data.

Thank You for Listening!

Mail:yusuke_s@m.sys.t.u-tokyo.ac.jp

How to Calculate Cramer's V

Components①/②	X	Y	Sum
A	30	10	40
B	20	10	30
C	10	10	30
Sum	60	40	100

Original Data

Components①/②	X	Y	Sum
A	24 (100*0.6*0.4)	16 (100*0.6*0.4)	40(40%)
B	18	12	30(30%)
C	18	12	30(30%)
Sum	60(60%)	40(40%)	100

Expected Frequency

Components①/②	X	Y
A	$(30 - 24)^2/24$	$(10 - 16)^2/16$
B	$(20 - 18)^2/18$	$(10 - 12)^2/12$
C	$(10- 18)^2/18$	$(10 - 12)^2/12$

Calculate
(Original Data – Expected Frequency)² / Expected Frequency
at each combination

The sum is χ^2 (In this case, 13.2)

And Cramer's V is Calculated by this formula:

$$v = \sqrt{\frac{\chi^2}{n(k-1)}}$$

n: number of samples, k: smaller number of kinds of two elements

The Methods of IGTA+ Algorithm

IGTA

1. Each element is initially placed in its own cluster.

2. Calculate the Coordination Cost of the Cluster Matrix by the following formulas:

$$IntraClusterCost=(DSM(j,k)+DSM(k,j))*ClusterSizepowccm$$
$$ExtraClusterCost=(DSM(j,k)+DSM(k,j))*DSMSizepowcc$$
$$TotalCost=\sum IntraClusterCost+\sum ExtraClusterCo$$

3. Randomly chooser an element.

4. Calculate bid from all clusters for the selected element by the following formula:

$$ClusterBidj=inoutpowdepClusterSizepowbidj$$

5. Randomly choose an integer in [0, *rand_bid*).

6. Calculate the total Coordination Cost if the selected element becomes a member if the cluster with the highest bid (use the second highest bid if 1 is choosen in step 5.

7. Randomly choose an integer in [0, *rand_accept*).

8. If new Coordination Cost < the old Coordination Cost or 1 is chosen in step 7, make the change.

9. Go back to step 3 until repeated a set number of times.

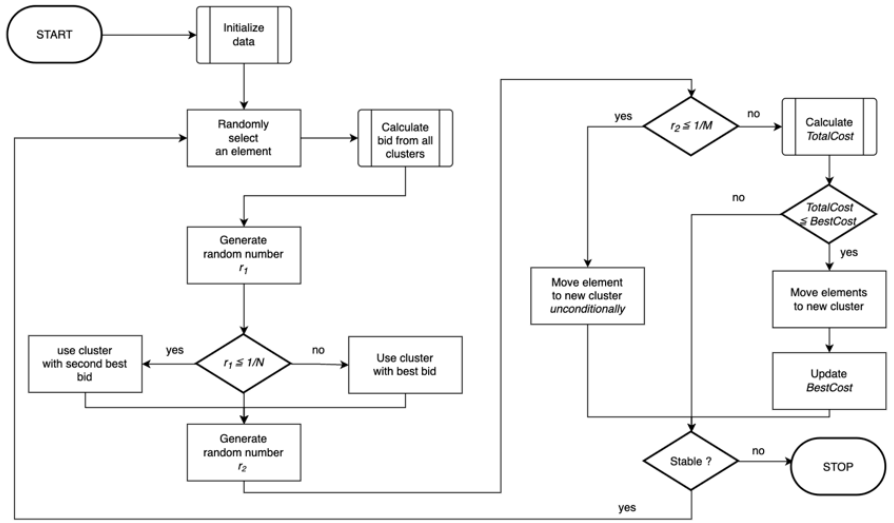
IGTA-Plus

- Suppressing Multicluster Allocation (SMA)

Randomly assign the selected element to one of the clusters with the highest bid.
Calculate the Coordination Cost more efficiently by the following formulas:

$$IOIO=ClusterMatrixClusterMatrix\cdot DSMDSM\cdot ClusterMatrixTClusterMatrixT$$
$$IOIntra=\sum k=1nIOIOkk\cdot ClusterSizepowcck$$
$$IOExtra=(\sum i=1n\sum j=1nIOIOij-\sum k=1nIOIOkk)\cdot DSMSizepowcc$$
$$TotalCost=IOIntra+IOExtra$$

- Improved termination criterion
  Keep a list of elements that have not yet been tried. When the list is empty, the algorithm terminates. Reset the list when move is made.



Borjesson, F, & Hölttä-Otto, K. "Improved Clustering Algorithm for Design Structure Matrix." *Proceedings of the ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 3, 38th Design Automation Conference, Parts A and B. 2012, pp. 921-930

Example of Calculation Results of Cramer’s V

- The Cramer's V is high if one element is changed when the other is also changed.
- Focusing on areas with high Cramer’s V, it is possible to identify components that are being changed at the same time.

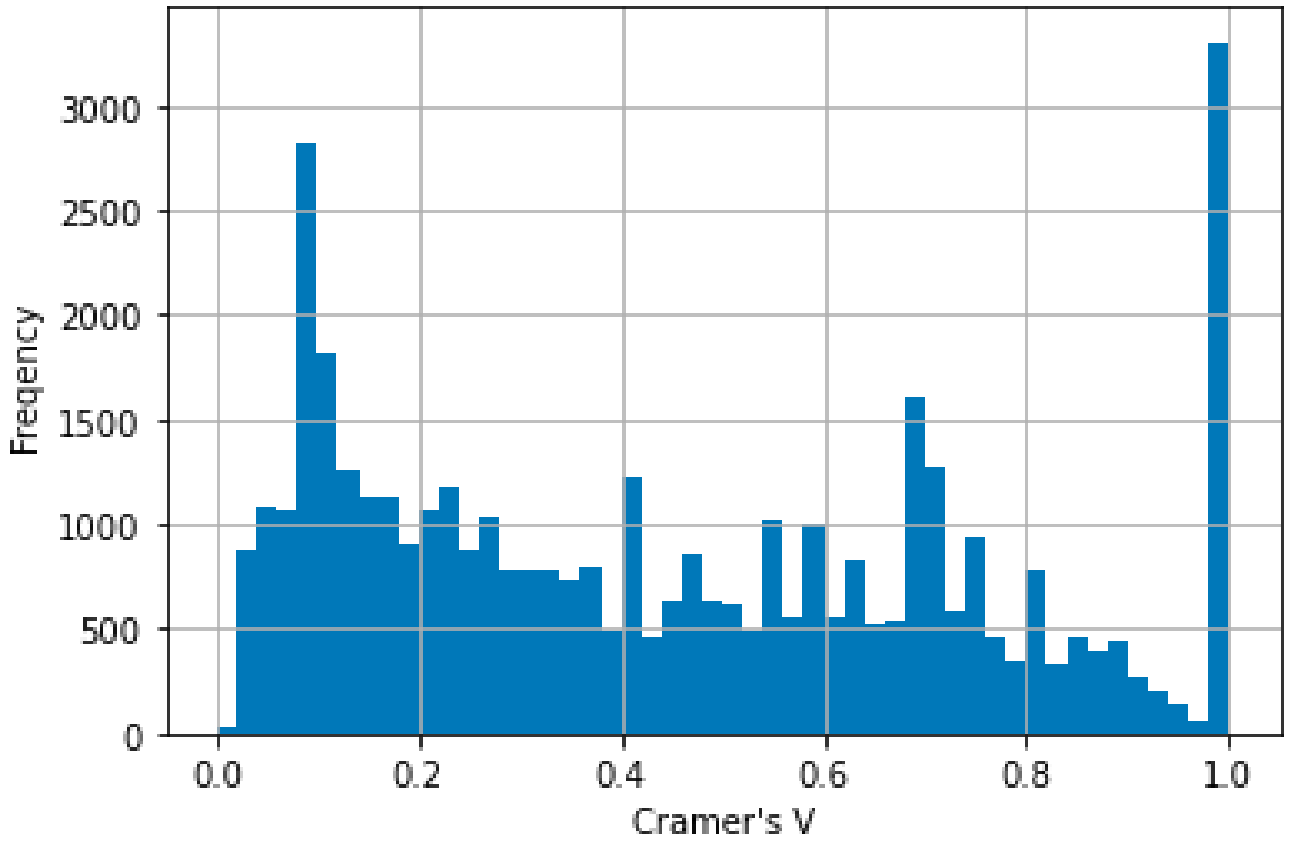
	M+Main Engine	M+Turbocharger	M+Main Generator Engine	M+Main Generator
M+Jacket Cooling Fresh Water Cooler	1	0.438986	0.375849	0.375849
M+Low Temp. Cooling F.W. Cooler	1	0.612372	0.5	0.5
M+D/G Low Temp. Cooling F.W. Cooler	0.707107	0.664057	0.818579	0.818579
M+Main Lubricating Oil Cooler	0.892354	0.732828	0.517693	0.596243
M+Central Cooling Fresh Water Cooler	0.750309	0.547473	0.439389	0.439389
M+Urea Cooler	1	0.612372	0.5	0.5
M+Main Engine M.G.O.Cooler	0.218218	0.218218	0.313392	0.773443

Example of Incorrect Information Made by Cramer's V

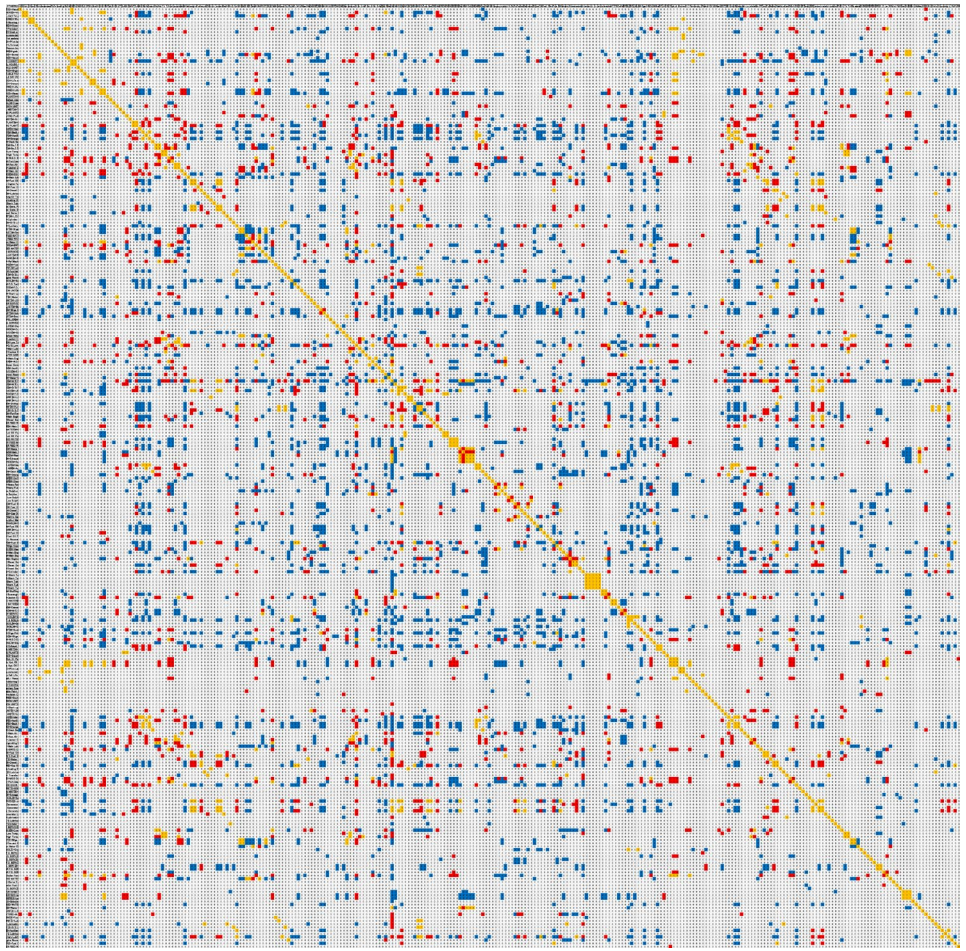
- There is no design coordination relationship between the cooling system of a diesel generator and a lifeboat, and they are not changed at the same time in practice.
- However, when one specification is changed, if the content of the other specification is also changed, the Cramer's V will have a high value.

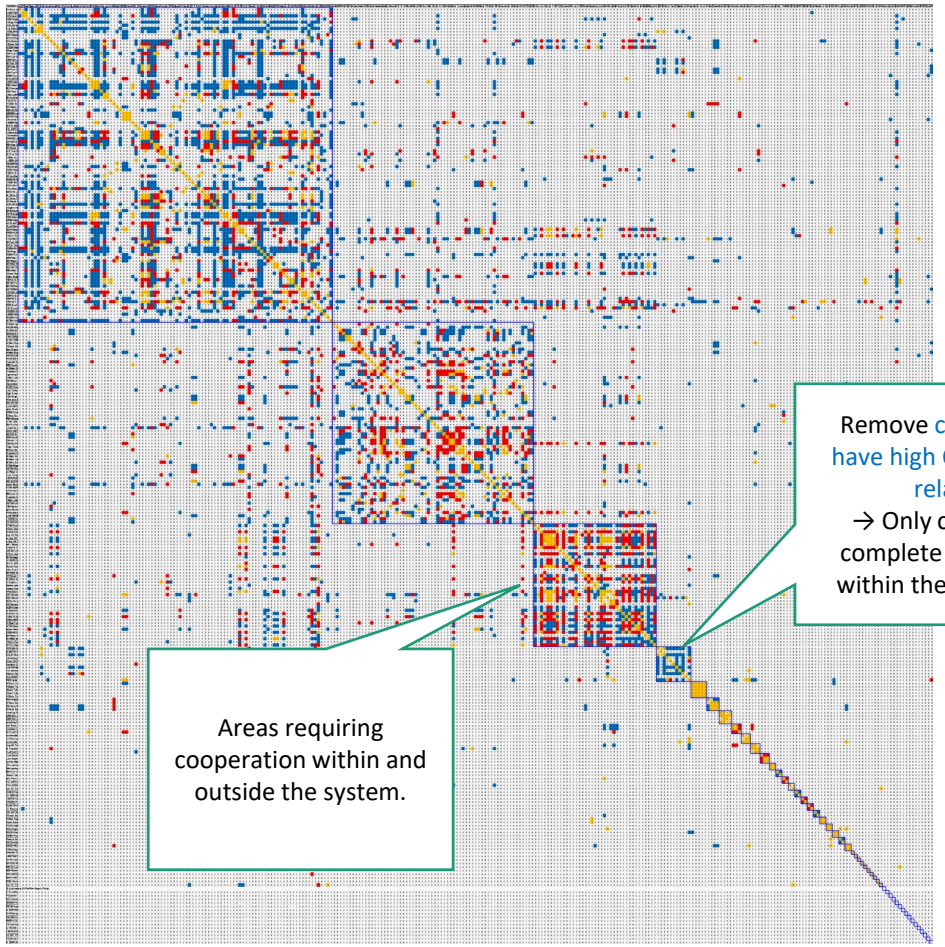
	H+Life boat	H+Rescue boat
M+Main Air Compressor	1	1
M+Back up Air Compressor	1	1
M+Auxiliary Air Compressor	0.816497	0.816497
M+Emergency Air Compressor	1	1

Frequency of Cramer’s V



DSM Clustering after System-level Check





The relationship between systems and accuracy

- Components with high Cramer V (and are likely to change at the same time) and the same lineage tend to be related by design.

related for design?	System	Num
Yes	Same	157
	Different	94
No	Same	241
	Different	3118

The number of pairs of components that have high Cramer's V

39% correct response rate for components with high Cramer V and the same lineage.

Can the accuracy be further improved by referring to indicators of network analysis, etc., after narrowing down to the same lineage?

Can indicators for determining correctness and incorrectness of the relationship between components belonging to different lineages be identified and manual work reduced?