9th European Workshop on Structural Health Monitoring
July 10-13, 2018, Manchester, United Kingdom

Hierarchical Fisher-information-matrix-based Clustering

Saeid Allahdadian[1], Michael Döhler[2], Carlos Ventura[3] and Laurent Mevel[4]
1 University of British Columbia, Canada, asaeid@alumni.ubc.ca
2 Inria, I4S, Campus de Beaulieu, France, michael.doehler@inria.fr
3 University of British Columbia, Canada, ventura@civil.ubc.ca
4 Inria, I4S, Campus de Beaulieu, France, laurent.mevel@inria.fr

Abstract

In this paper, a clustering approach, namely Hierarchical Fisher-information-based Clustering (HFC), is proposed for clustering the similar elements in a structure, based on their effect on the modal parameters and dynamic behaviour of the structure. This clustering approach is indispensable and valuable for any damage localization approach, as always in practice, the SHM systems include very low number of sensors compared to the number of elements of a structure (or degrees of freedom). Therefore, a clustering approach is a great tool in assessing the possible localization resolution. Furthermore, a clustering approach is necessary to be used for one of the powerful tests, i.e. MinMax test, in Statistical Subspace Damage Localization (SSDL) method. The robustness of the SSDL method on localizing damage in real structures was demonstrated in the literature. In here, the HFC clustering approach along with its effects on the damage localization results for a real test structure, the Yellow Frame, will be presented. It will be shown that the HFC approach can robustly cluster the similar elements of a structure compared to a popular and well-known clustering method, i.e. *k*-means. The results will be shown and compared.

1. Introduction

A significant component of structural health monitoring (SHM) is damage localization methods which process the data with the purpose of localizing damages in the structures after detecting them. One of these methods with a proven robustness in detecting and localizing the damages in real structures is the Statistical Subspace Damage Localization (SSDL) method (1,2).

There are numerous researches found in the literature dealing with damage identification of structures. These methods use different responses of structure and are reviewed in some review papers such as (3,4). One category of these techniques includes the methods employing statistical tests in identifying the damage. The SSDL method is categorized in this group.

Since the damage in a structure results in changes in its natural frequencies and modeshapes, monitoring these modal parameters can be used in identifying the damage. However, identification of modal parameters is not usually accurate (especially for higher modeshapes) and it needs manual processing of the data; therefore they are not appropriate for real-time health monitoring. In the SSDL method, there is no need to estimate the natural frequencies and modeshapes, making this approach capable of being used in real-time monitoring of structures. In this way, the whole eigenstructure, i.e. modeshapes and natural frequencies, of the measurements are included in the damage detection and the focus is not only on dominant frequencies. Including higher modes in this evaluation makes the damage detection approach more robust, considering that the main effect of local damages is on higher modeshapes.

The SSDL method was shown to be able to robustly detect the damages (5,6) and localize them in real structures (6,7).

In this approach, i.e. SSDL, in addition to the data driven residual (8), an analytical model of the structure, e.g. finite element model, is used to localize the damage (9). This analytical model needs to be created for the system in the reference state (healthy structure with no damage) only and there is no need to update the model for the damage testing. This model is used to calculate Jacobians of the data driven residual to each physical parameter of the structure, e.g. members stiffness or sectional properties. The Jacobians are a bridge in connecting the data driven domain, i.e. residual, to the analytical model domain, i.e. physical parameters. For this purpose, an important link between the data-driven residual and the finite element model is needed, which is made through a sensible parameter choice by clustering and a respective sensitivity computation. In this paper, a novel clustering approach, namely Hierarchical Fisher-information-matrix-based Clustering (HFC), is described that by an applicable scheme can be used for choosing the parameters for the required sensitivity computation in realistic applications (6).

This clustering approach, i.e. HFC, is indispensable and valuable for any damage localization approach, as always in practice the SHM systems include very low number of sensors compared to the number of elements of a structure. For this purpose, clustering the elements which are dynamically similar to each other (we call them *close* elements) makes the approach more accurate and the results will be more reliable with an understanding of the amount of resolution of detection we are able to gain.

In this paper, the SSDL method will be presented and then the HFC approach will be described by theory and subsequently, the results on a real test structure, will be demonstrated in practice. The SSDL formulations in section 2 will be modified and adjusted based on the clusters in section 3. Finally, the results on the real test structure, i.e. the Yellow Frame (6,10), will be shown and discussed.

2. Statistical Subspace Damage Localization (SSDL) method

As shown in (1,9), the SSDD framework offers the possibility to detect parameters that are responsible for the changes in the residual function. Hence, by defining the parameter set as the collection of parameters linked to a finite element model, we can detect the parts of the model that has changed due to this damage. In this approach in addition to the data driven residual, an analytical model of the structure, e.g. finite element model, is used to localize the damage.

Firstly, the state-space representation of the dynamic model of a structure can be written in discrete-time space as

$$\begin{cases} x_{k+1} = Fx_k + w_k \\ y_k = Hx_k + \varepsilon_k \end{cases} \quad \textbf{(1)}$$

where x_k is representing the state of the system in step k and y is the measured output (e.g. from sensors). The F and H matrix are, in order, the state transition matrix and observation matrix. In a general form, the noise is imposed in the model on the state as w_k and on the measured output as ε_k. The importance and effect of these noise parameters on the robustness of the SSDL method were previously investigated by the authors in (6,11,13,14).

By the use of output-only covariance based subspace system identification method (15), a residual vector will be defined. For this, define the Hankel matrix $\mathbf{H}_{p+1,q}$ as

$$\mathbf{H}_{p+1,q} = \begin{bmatrix} R_1 & R_2 & \cdots & R_q \\ R_2 & R_3 & \cdots & R_{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ R_{p+1} & R_{p+2} & \cdots & R_{p+q} \end{bmatrix} = \text{Hank}(R_i) \,. \tag{2}$$

Since, this Hankel matrix includes the dynamic properties of the system, we can monitor the changes (damages) in the system by comparing this matrix in two conditions of the structure: the undamaged (reference state) and test state (damaged or undamaged). This comparison can be done by using the following characterization of the Hankel matrix and its left null space in reference state, i.e. S_0, as

$$S_0^T \mathbf{H}_{p+1,q} = 0 \,. \tag{3}$$

By computing the left null space matrix in the reference state and recomposing the estimate of the Hankel matrix in test state as $\hat{\mathbf{H}}_{p+1,q}$, and in view of (3), the residual ζ can be defined as

$$\zeta = \sqrt{N}\ \text{vec}(S_0^T \hat{\mathbf{H}}_{p+1,q}) \tag{4}$$

where N is the number of samples. It can be shown that this residual is normally distributed in each state, and therefore, the following hypothesis can be defined as

$$\zeta \rightarrow \begin{cases} \mathcal{N}(0,\Sigma) & \text{under } H_0 \\ \mathcal{N}(J\delta\theta,\Sigma) & \text{under } H_1 \end{cases} \tag{5}$$

in which Σ is the asymptotic covariance, $\delta\theta$ can be a vector of changes (damages) in elements and J is the asymptotic sensitivity of the residual computed with respect to the physical parameters θ.

2.1 Parametric Hypothesis Test

In order to test hypothesis (5), a generalized likelihood ratio (GLR) test is used (16). the GLR test is asymptotically χ^2-distributed with degrees of freedom equal to $d{=}\text{rank}(J) = \dim(\theta)$ and non-centrality parameter $\delta\theta^T J^T \Sigma^{-1} J \delta\theta$ under H_1 and 0 under H_0, and thus

$$\chi^2 = \zeta^T \Sigma^{-1} J \left(J^T \Sigma^{-1} J \right)^{-1} J^T \Sigma^{-1} \zeta \,. \tag{6}$$

The χ^2-variable is the parametric representation of a damage index and is compared with a threshold of safety. Since its distribution is shifted with the given non-centrality parameter under H_1, if its value surpasses this threshold, it shows that the condition of the structure is being changed. Hence, it indicates that a damage in the system has happened.

2.1.1 Sensitivity based approach
The damage in the structure can be modelled as a change in a parameter, e.g. p_k, of the analytical model as used in (1,8). Now, define

$$\tilde{\zeta} = \hat{\Sigma}^{-\frac{1}{2}}\zeta \qquad \text{and} \qquad \tilde{J} = \hat{\Sigma}^{-\frac{1}{2}}\hat{J} \tag{7}$$

where $\hat{J}$ is the consistent estimate of J and $\hat{\Sigma}^{-\frac{1}{2}}$ is the matrix square root of the inverse of the estimate of $\hat{\Sigma}$. Such decomposition of the covariance matrix to its inverse roots is possible since it is positive definite. Based on (7) and in view of (5), the new residual is distributed as

$$\tilde{\zeta} \rightarrow \begin{cases} \mathcal{N}(0,\mathbf{I}) & \text{under } H_0 \\ \mathcal{N}(\tilde{J}\delta\theta,\mathbf{I}) & \text{under } H_1 \end{cases}. \tag{8}$$

Hence, the GLR test (6) can be shown by the χ^2-test for each parameter p_k (a physical parameter) of the structure as

$$\chi^2(p_k) = \frac{\tilde{\zeta}^T \tilde{J}_k \tilde{J}_k^T \tilde{\zeta}}{\tilde{J}_k^T \tilde{J}_k} \tag{9}$$

where $\tilde{J}$ is a matrix collecting all vectors $\tilde{J}_k$ for $k = 1 \cdots N_\rho$, and N_ρ is the total number of parameters. Same as (6), if element p_k would be damaged, (9) will be increased and surpass a safety threshold.

2.1.2 MinMax test
In this test, the effect of changes in other elements on the χ^2-test value of an element is removed. This removal is achieved in (9,17) by projecting the residual on the element being tested and removing the projections from other elements (being "blind" to other elements). Therefore, in this test the computed χ^2-value for an element conveys only the information from the change in that element while being blind to the changes in other elements. This will reduce the chance of false positive result for the undamaged elements, while the sensitivity-based approach is prone to it.
The corresponding robust χ^2-test, i.e. χ^{*2}, can be defined based on (9) as

$$\chi^{*2}(p_k) = \zeta_a^{*T} F_a^{*-1} \zeta_a^* \tag{10}$$

where ζ_a^* is a robust residual corresponding to element *a* (to be tested) with physical property p_k and the F_a^* is the corresponding robust Fisher information matrix. The definition and implementation of these parameters can be seen in (6,9). It should be noted that although for a damaged element the robust χ^2-test has smaller or equal value than the sensitivity based approach, but the effect of damage in other elements is removed from this factor. This makes the damaged elements more distinguishable than the other test, i.e. sensitivity based approach.

3. Hierarchical Fisher-information-matrix-based Clustering (HFC)

The number of physical parameters in a structure is usually higher than the identified modal parameters and hence, the sensitivity (Jacobian) matrix J is usually a "wide" matrix. Since the number of sensors is considerably less than the number of DOFs of the structure, the resolution of the identification in terms of elements is not high and the columns of the Jacobian corresponding to *close elements* are pointing to the same direction. This causes the χ^2-test of the *close* elements to react the same way. This *closeness* stems from the modal behaviour of the elements which in turn is related to their geometrical and physical closeness and modal direction in the considered modeshapes.

Furthermore, the *close* elements cannot be directly treated in the MinMax test, because the Jacobian matrix is required to be full column rank. The reason is that the MinMax test insures of seeing purely the change in the tested element by removing the other elements effect. However, if two elements are *close* and one of them is damaged, when testing the damaged element the effect of damage is removed from the test by its *close* element. This will reduce drastically the χ^2-test reaction to damage for the damaged element and therefore generates false negative results. In order to remove this effect, clustering of elements is necessary.

The closeness of the elements can be identified from the directions of their corresponding Jacobian vectors which will be used in the clustering procedure. Figure 1 illustrates how the vectors of Jacobians of *close* elements look like.

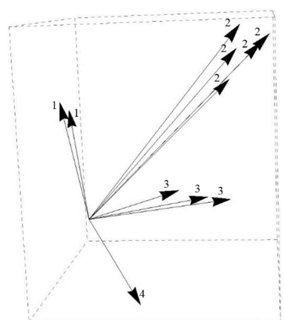

Figure 1. Schematic illustration of closeness of Jacobian vectors

In clustering the columns of Jacobian matrix, the normalized Jacobians are used because in both the sensitivity approach and MinMax test, $\tilde{J}$ is the basis of χ^2-test. This will assure that the Jacobian vectors are clustered consistent to the χ^2-test in (9) and (10) as the directions of columns in $\tilde{J}$ are not necessarily the same as J.

Since the scaling of the columns of Jacobian will not affect the test value the angles of them in the vector space is the only parameter for measuring *closeness*. Therefore, the Columns of Jacobian should be normalized to unit vector prior to clustering, to remove any effect of their scaling on the clustering approach. This normalization can be performed as

$$\bar{J}_k = \frac{\tilde{J}_k}{\left\| \tilde{J}_k \right\|} \tag{11}$$

where $\bar{J}_i$ is the normalized Jacobian column corresponding to element k with unit length.

It should be noted that the final χ^2-value of all the tests proposed herein, is not affected by the scale (size) of Jacobians. Therefore, one way in dealing with the elements with small size of Jacobian columns $\tilde{J}_k$ is to remove them before the testing. The reason of

this removal is the fact that, with similar variances, the columns of Jacobian matrix with small scaling will pose higher error on the test than the larger ones. However, the decision of removal or keeping these vectors are based on the engineering judgment of the tester.
Two approaches of clustering *close elements* are described in the following.

3.1 k-means clustering

The k-means clustering is a vector quantization approach frequently used in signal processing, image processing and machine learning fields. In this algorithm, firstly k number of groups is assumed and then randomly k points (vectors) in the space are selected from the total N_ρ points as the centroids for these groups. Subsequently, the other points in the space are categorized to each of these points based on their minimum distance to the centroids. Iteratively, the mean of each group is calculated and each point in the space is re-associated to the group with closest centroid. This iteration converges when no point is re-associated to other groups. The clusters and their centroids are illustrated schematically in the following figure:

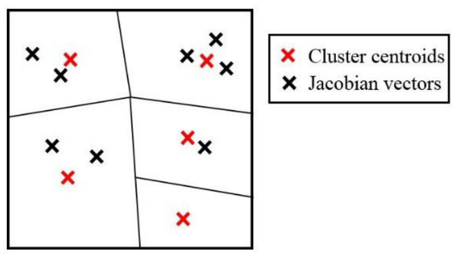

Figure 2. Schematic illustration of *k*-means clustering

Although this algorithm is frequently used in clustering approaches in literature, there are some important shortcomings of it which will be described in here.
This algorithm is highly dependent on the number of groups, i.e. k, and the starting random points. It is not guaranteed to converge while it only can converge to local minima. Therefore, different starting points can result in different classifications. Moreover, the number of groups of the structural elements are unknown and the resultant χ^2-test is highly depending on that. Furthermore, even after convergence of this algorithm *close* elements are not necessarily categorized into the same group and hence the results are not promising from the Min-Max test. This algorithm is also computationally not efficient since the running time of it is given as $O(N_\rho ki)$, where i represents the number of iterations to convergence.
Due to these disadvantages of this approach, the Fisher information matrix is used in clustering of the elements as proposed in the following subsection.

3.2 HFC

The *closeness* of the elements of the Jacobian matrix $\bar{J}$ can be assessed by the correlation between the vectors. This correlation is calculated as the normalized Fisher information matrix f ,

$$f = \bar{J}^T \bar{J} \ . \qquad \textbf{(12)}$$

Each element f_{ij} of this matrix corresponds to the *closeness* of vectors i and j of the normalized Jacobians and is computed as

$$f_{ii} = 1, \;\; f_{ij} = \frac{\tilde{J}_i^T \tilde{J}_j}{\left\| \tilde{J}_i \right\| \left\| \tilde{J}_j \right\|} \ . \qquad \textbf{(13)}$$

The normalized Fisher information matrix (NFIM) is positive definite and symmetric due to its composition in (12). An element of this matrix with value near 1 corresponds to the *close* vectors of the Jacobians and a small value near 0 shows the opposite. Therefore, the clustering can be done by grouping the elements corresponding to high values in the normalized Fisher information matrix (NFIM). For this purpose, a hierarchical clustering approach is used to group the elements based on upper triangle of NFIM. Figure 3 is a dendrogram depicting the hierarchical clustering of 32 elements based on their corresponding values in *f*. In this picture, as the height of the connections are increased, their *closeness* is decreased; the height of the connections is an inverse ratio of their corresponding value in the NFIM.

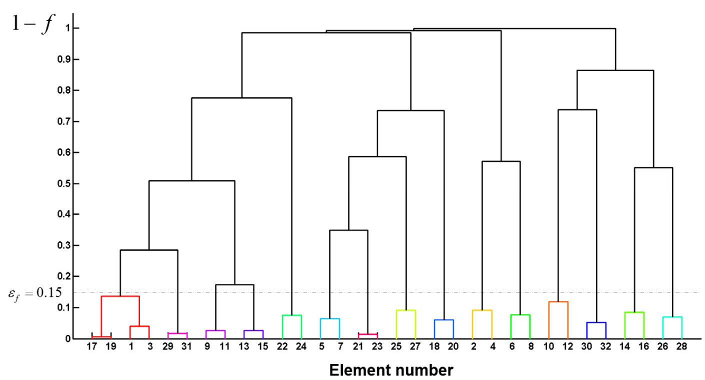

Figure 3. Hierarchical clustering of 32 elements

The elements clustered at the lowest level, two by two in Figure 3 are not distinguishable in terms of damage from each other and hence they become clustered in the first step. This happens when two elements are very *close*.

After having the classification shown in Figure 3, a threshold ε_f needs to be selected on the difference of f_{ij} from 1. This threshold defines the amount of *closeness* of vectors needed in order to classify them as one vector. The dashed line in Figure 3 shows the threshold $\varepsilon_f = 0.15$ from which the elements are clustered into 15 clusters.

By increasing this value, the number of clusters will decrease, the resolution of the damage localization is decreased and the uniqueness (perpendicularity) of the clusters will increase. Therefore, there is a compromise between the resolution of damage localization and uniqueness of clusters which can be adjusted by ε_f . The optimal ε_f can be chosen by minimizing it while having the constraint of sufficient perpendicularity of the clusters. This is achieved by looking at the dendrogram of the clustering and the resultant NFIM in an iterative manner, only in the reference state.

After having this clustering, the mean of the vectors associated to each cluster is used as the centroid of that cluster. The NFIM of the 32 elements before and after clustering are compared in Figure 4.

It can be seen from this figure that the number of elements with high values (warm colors) of f_{ij} are reduced after clustering. To further reduce the remaining orange spots, we need to increase ε_f which as discussed will reduce the resolution of the damage localization.

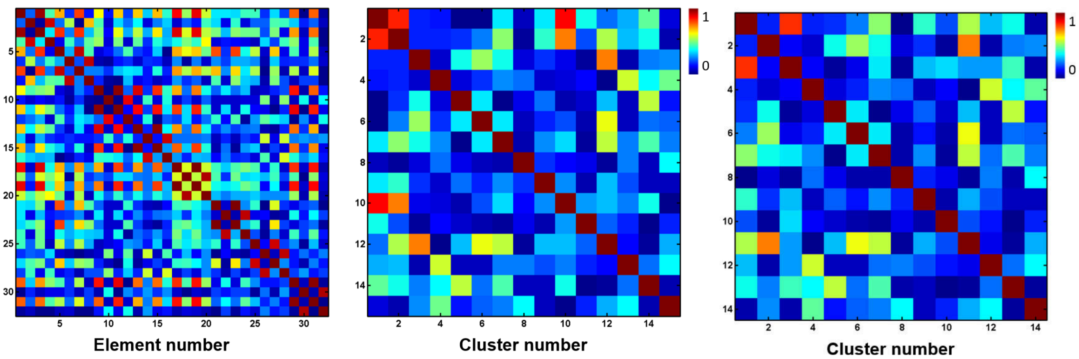

Figure 4. Normalized Fisher information matrix (NFIM) of 32 elements before clustering (left); after clustering with $\varepsilon_f = 0.15$ into 15 clusters (middle) and after clustering with $\varepsilon_f = 0.23$ into 14 clusters (right)

3.3 Application of clusters in tests

After clustering the Jacobians matrices, the clusters need to be applied in the sensitivity based and MinMax test. This application can be done in different ways for each test, which will be described in here.

3.3.1 Clusters in Sensitivity approach test

By clustering the normalized Jacobian matrix $\bar{J}$, into N_c clusters, the sensitivity test (9) can be performed on the centroids of the clusters from k-means or HFC approach as

$$\chi^2(p_k) = \frac{\tilde{\zeta}^T C_j C_j^T \tilde{\zeta}}{C_j^T C_j}, \quad \text{for} \quad \forall k \in \mathbf{C}_j,\ j = 1 \cdots N_c. \qquad \textbf{(14)}$$

In (14), $\mathbf{C}_j$ is the jth cluster with centroid C_j. In view of (14), all the elements inside a cluster will be identified as damaged or undamaged based on the χ^2-value of their corresponding cluster centroid.

3.3.2 Clusters in MinMax test

The use of clusters in the MinMax test can be performed in two approaches based on the definition of $\tilde{J}_a$.

First approach: In this approach, $\tilde{J}_a$ is chosen as the centroid of one cluster to be tested and $\tilde{J}_b$ is formed consistently from the centroids of other clusters. Therefore,

$$\begin{cases} C_a = C_j \\ C_b = \left[C_1 \cdots C_{j-1} C_{j+1} \cdots C_{N_c} \right] \end{cases}, \quad j = 1 \cdots N_c. \qquad \textbf{(15)}$$

Thus, all the elements in the cluster are treated similarly as damaged or not damaged.
***Second approach*:** In this approach each of the elements in a cluster can be tested which results in a higher resolution than the first approach. To achieve that, the Jacobian is discretized as

$$\begin{cases} C_a = \bar{J}_k \\ C_b = \left[C_1 \cdots C_{j-1} C_{j+1} \cdots C_{N_c} \right] \end{cases}, \; k \in \mathbf{C}_j, \; j = 1 \cdots N_c . \quad \textbf{(16)}$$

Therefore, each element in a cluster is tested while removing the effects of other clusters. Since, the other clusters do not include the elements *close* to the element being tested, the damage effect will not be removed from the resulting χ^2-test.
After partitioning the Jacobians from either approach the normalized Fisher information matrix F_c can be defined consistently from the cluster centroids as $F_C = \left[C_a \quad C_b \right]^T \left[C_a \quad C_b \right]$ and partitioned similar to (15) as

$$F_C = \begin{bmatrix} F_{C_{aa}} & F_{C_{ab}} \\ F_{C_{ba}} & F_{C_{bb}} \end{bmatrix} . \quad \textbf{(17)}$$

By defining the residuals as $\tilde{\zeta}_{C_a} = C_a^T \tilde{\zeta}$ and $\tilde{\zeta}_{C_b} = C_b^T \tilde{\zeta}$ the robust residual will be $\zeta^*_{C_a} = \tilde{\zeta}_{C_a} - F_{C_{ab}} F_{C_{bb}}^{-1} \tilde{\zeta}_{C_b}$. Hence, the robust χ^2-test writes as

$$\chi^{*2} = \zeta^{*T}_{C_a} F^*_{C_a} \zeta^*_{C_a} \quad \textbf{(18)}$$

where $F^*_{C_a} = F_{C_{aa}} - F_{C_{ab}} F_{C_{bb}}^{-1} F_{C_{ba}}$.
Based on the definition of C_a from the two approaches, the χ^2-test (18) identifies the location of the damage in each cluster or element. It should be noted that although in the second approach the test is performed on each element, but the resolution of damage localization is highly depending on the number and location of sensors. Therefore, the resolution of the test is higher in the second approach while limited by the available information from data.
Based on the characteristics of the MinMax test and HFC clustering, the combination of these two with the second approach should be a very robust approach. From such combination, the detection resolution is element basis and the effect of each element is investigated without the interruption of condition of other elements.
The other alternative is the sensitivity based approach with/out clustering. There is no need to cluster this method. From this approach the resolution is element based if no clustering is used. However, the condition of other elements will affect the result of testing a specific element which may lead to possible false positive results.

4. Case study: the Yellow Frame

In order to demonstrate the proposed HFC approach the result acquired from a test performed on a real structure, namely the Yellow Frame (6) will be considered. The Yellow frame is a modular 4 story, scaled (1/3) steel frame established at the University of British Columbia (UBC). Several damage scenarios are designed and tested by removal of braces of the structure from which one of them is demonstrated by using the proposed HFC and *k*-means approach. This structure is 3.6 m high and is composed of 2 spans in each direction with the total length of 2.5 m. Each floor of the structure is carrying dead loads applied to the structure by using 4 steel plates distributed on each level. This structure was used in evaluating several other damage detection techniques as well such as (18,19).

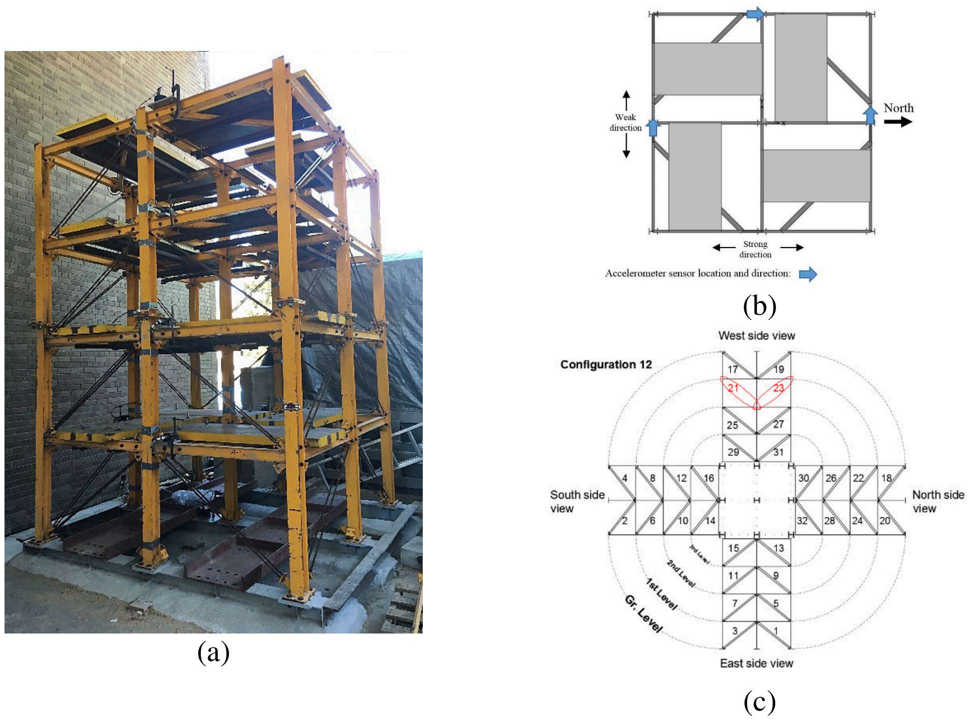

Figure 5. (a) The Yellow frame structure, (b) the schematic plan of the structure showing the location of sensors, (c) the numbering of the braces of the structure

The FE model of this structure is built and the sensitivity analysis of the mode shapes and natural frequencies with respect to each brace is computed using a finite difference approach. It should be noted that the sensitivities are all calculated from the analytical model results, since as it was shown in (6,7) it would yield better results.

4.1 Clustering

The elements of the Yellow frame need to be clustered before damage localization. It should be noted that the braces located in the same story level and in the same side of the structure are *close* as can be seen in Figure 5. Therefore, ideally the clustering methods should consider them in the same group. Since there are 16 pairs of these close braces present in the structure, the number of the resultant clusters should be less than 16. This can be used in checking the output of the clustering approaches.

4.1.1 Hierarchical Fisher-information-matrix-based clustering

The clustering for the structure by HFC method with using threshold of $\varepsilon_f = 0.15$ is shown in Figure 6Figure .

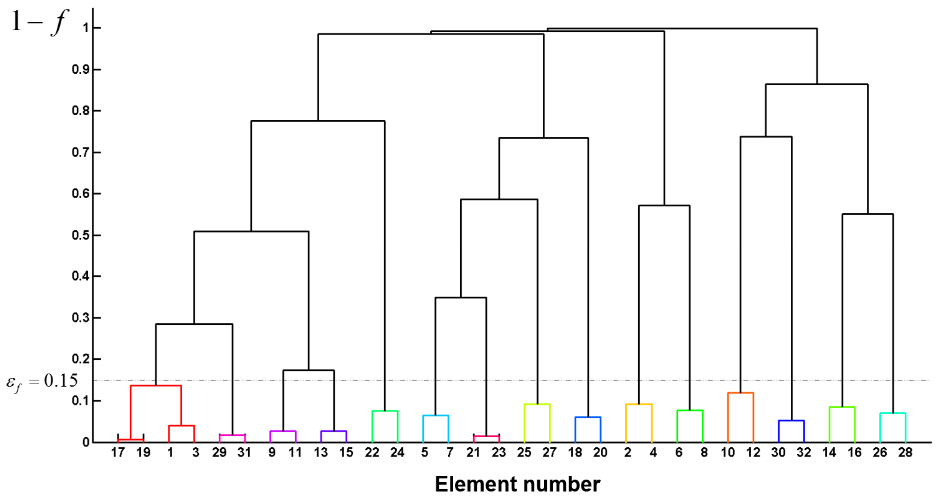

Figure 6. Dendrogram depicting the Hierarchical Fisher-information-matrix-based clustering (HFC)

It is observed in Figure 6, that Jacobians resulted in clustering scheme with 15 clusters which is less than 16 clusters. Furthermore, all the braces expected to be *close*, are in the same group, in view of Figure 6. Using this clustering scheme, the NFIM is evaluated and shown in Figure 7 for clustered and unclustered Jacobians.

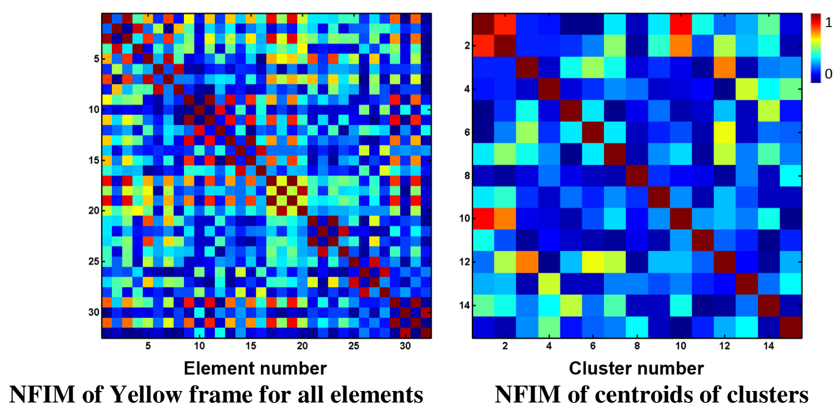

Figure 7. Normalized Fisher information matrix (NFIM) for the HFC-clustered and unclustered Jacobians

By the use of these clusters, the damaged structure is being measured and the residual values are calculated for each element. The residual is cacluclated using different methodologies, i.e. sensitivity-based approach and MinMax test. For the MinMax test, the test is performed both with the clustered and unclustered Jacobians to demonstrate the necessity and effect of the clustering on this approach.

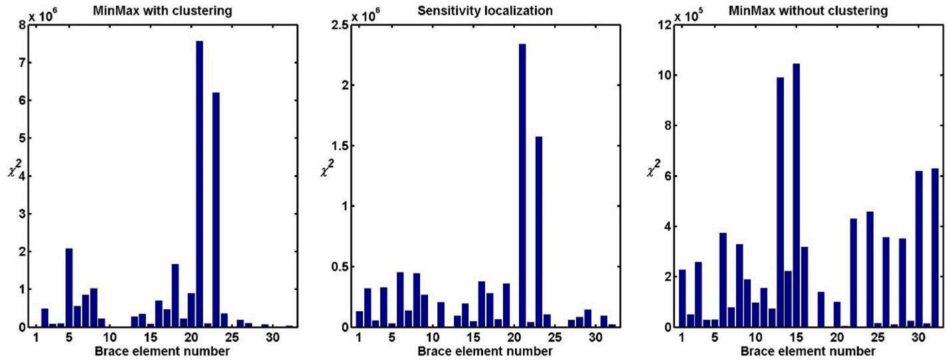

Figure 8. Damage localization of the Yellow frame with HFC clustered Jacobians

It is seen that both damaged elements (damaged brace/s) can be identified with acceptable accuracy.
It should be noted that the *close* braces, i.e. every couple of braces in each level at each side, cannot be distinguished in terms of being or not being damaged from each other. In other words, if one of the braces is damaged, the other *close* brace also reacts in the test as to be damaged.
The MinMax method with the HFC clustering, can clarify (magnify) the damage compared to the sensitivity based damage localization. Furthermore, the MinMax method without clustering is incapable of localizing the damage as predicted.

4.1.2 k-means clustering

In order to show the performance of *k*-means in clustering the elements, the *k*-means approach is used in clustering the Jacobians. As mentioned before, the results are not unique and one of the clustering schemes is presented in Figure 9. The number of clusters is chosen as 15 to be comparable to the results from HFC approach. Again it should be noted that the objective function of *k*-means is chosen as cosine between input vectors, i.e. columns of Jacobian. The results were also observed to be the same when choosing the correlation between vectors as the objective function.
It can be seen in Figure 9, that some of the *close* braces are not in the same cluster. Therefore, in the NFIMs corresponding to these clustering schemes, several red color spots are existing which are related to these close elements. Based on these clustering schemes, it is expected that if the damage happens in close elements that are not in the same cluster, e.g. elements 21 and 23, the damage would not be identified.

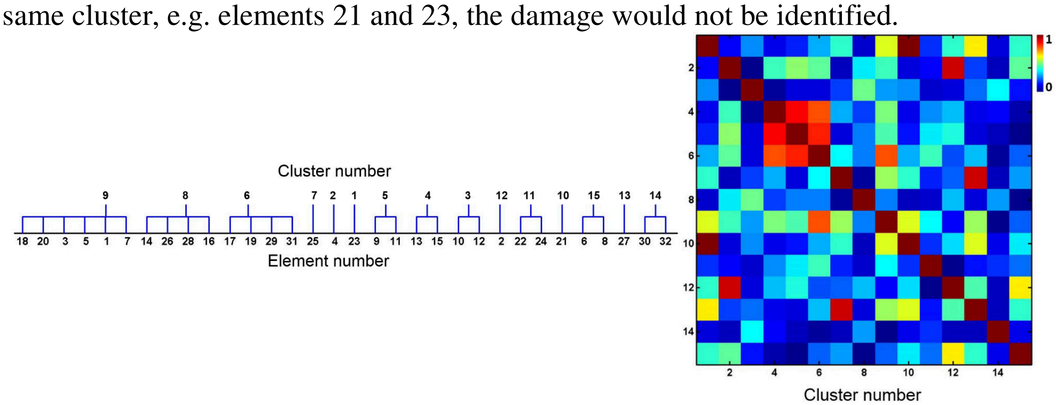

Figure 9. Clustering acquired with *k*-means approach, left dendrogram and right is NFIM

In order to assess the clustering scheme computed by *k*-means approach, in here, the SSDL test is performed to demonstrate the effect of *close* elements that are not clustered properly on the MinMax method. As can be seen in Figure 10, the damage in *close* elements not clustered properly is not detected.

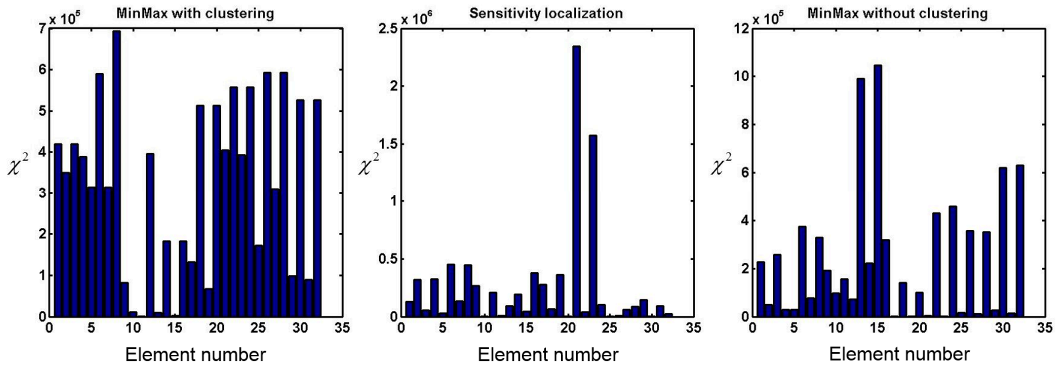

Figure 10. Damage localization of the Yellow frame with *k*-means clustered Jacobians

Because elements 21 and 23 are not in the same cluster they cannot be identified using MinMax approach, as shown in Figure 10. Therefore, the clustering scheme acquired from the *k*-means approach are not appropriate in damage localization of this damage configuration. Subsequently, since there is no prior knowledge of the location of damage and closeness of elements in practice to check the *k*-means output and considering the unstable inherent of the *k*-means approach, it is not an appropriate clustering method to be used in SSDL method.

5. Conclusions

In this paper a clustering approach, namely HFC, was proposed and validated to cluster the elements which are similar in terms of their dynamic effect on the final modal parameters. This clustering is crucial in the use of MinMax test of SSDL technique. Also, this clustering can give us great information about the possible resolution of the damage localization method.

The HFC and *k*-means clustering were tested on a real test structure, the Yellow Frame. It was observed that HFC could cluster the *close* elements properly. The *k*-means approach was shown to be unstable and it could not cluster properly some *close* elements.

The MinMax and sensitivity based approach were used in localizing the damage. It was seen that the clustering scheme obtained from HFC is appropriate in localizing the damage with the MinMax test. Moreover, it was observed that the MinMax test with HFC can localize the damage in the structure as a multiple-damage scenario and had a better clarity on the damaged elements compared to the sensitivity based approach.

Hence, it was demonstrated that the HFC clustering approach is a stable and consistent method to cluster the elements robustly for the SSDL method, and the *k*-means approach is not a suitable method in clustering the elements for this method.

References

(1) E. Balmès, M. Basseville, L. Mevel, H. Nasser, and W. Zhou, "Statistical model-based damage localization: A combined subspace-based and substructuring approach," *Struct. Control Heal. Monit.*, vol. 15, no. 6, pp. 857–875, 2008.

(2) M. Döhler, L. Mevel, and Q. Zhang, "Fault detection, isolation and quantification from Gaussian residuals with application to structural damage diagnosis," *Annu. Rev. Control*, vol. 42, pp. 244–256, 2016.

(3) W. Fan and P. Qiao, "Vibration-based Damage Identification Methods: A Review and Comparative Study," *Struct. Heal. Monit.*, vol. 10, no. 1, pp. 83–111, Apr. 2010.

(4) S. W. Doebling, C. R. Farrar, and M. B. Prime, "A Summary Review of Vibration-Based Damage Identification Methods," *Shock Vib. Dig.*, vol. 30, no. 2, pp. 91–105, 1998.

(5) M. Döhler and L. Mevel, "Subspace-based fault detection robust to changes in the noise covariances," *Automatica*, vol. 49, no. 9, pp. 2734–2743, 2013.

(6) S. Allahdadian, "Robust Statistical Subspace-Based Damage Assessment," University of British Columbia, 2017.

(7) S. Allahdadian, M. Döhler, C. Ventura, and L. Mevel, "Damage localization of a real structure using the statistical subspace damage localization method," Sep. 2017.

(8) M. Basseville, L. Mevel, and M. Goursat, "Statistical model-based damage detection and localization: Subspace-based residuals and damage-to-noise sensitivity ratios," *J. Sound Vib.*, vol. 275, no. 3–5, pp. 769–794, Aug. 2004.

(9) M. Döhler, L. Mevel, and F. Hille, "Efficient computation of minmax tests for fault isolation and their application to structural damage localization," *IFAC Proc. Vol.*, 2014.

(10) S. Allahdadian, A. Mendler, C. Ventura, "The Yellow Frame – Experimental Studies and Remote Monitoring of the Structural Health Monitoring," in *IMACXXXVI, International Modal Analysis Conference, FL, USA Benchmark Structure*, 2018.

(11) S. Allahdadian, M. Döhler, C. E. Ventura, and L. Mevel, "On the Influence of Sample Length and Measurement Noise on the Stochastic Subspace Damage Detection Technique," Springer, Cham, 2016, pp. 35–46.

(12) S. Allahdadian, C. Ventura, P. Andersen, L. Mevel, and M. Döhler, "Investigation on the sensitivity of subspace based damage detection technique to damage and noise levels," in *IOMAC-International Operational Modal Analysis Conference*, 2015.

(13) S. Allahdadian, C. Ventura, P. Andersen, L. Mevel, and M. Döhler, "Subspace based damage detection technique: investigation on the effect of number of samples," in *CCEE-11th Canadian Conference on Earthquake Engineering*, 2015.

(14) S. Allahdadian, C. E. Ventura, P. Andersen, L. Mevel, and M. Döhler, "Sensitivity Evaluation of Subspace-Based Damage Detection Method to Different Types of Damage," Springer, Cham, 2015, pp. 11–18.

(15) P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*. Boston, MA: Springer US, 1996.

(16) M. Basseville, M. Abdelghani, and A. Benveniste, "Subspace-based fault

detection algorithms for vibration monitoring," *Automatica*, vol. 36, no. 1, pp. 101–109, 2000.

(17) M. Basseville, "Information criteria for residual generation and fault detection and isolation," *Automatica*, vol. 33, no. 5, pp. 783–803, May 1997.

(18) G. Bernagozzi, C. E. Ventura, S. Allahdadian, Y. Kaya, L. Landi, and P. P. Diotallevi, "Application of modal flexibility-based deflections for damage diagnosis of a steel frame structure," *Procedia Eng.*, vol. 199, pp. 2026–2033, Jan. 2017.

(19) M. D. H. Bhuyan *et al.*, "Transfer matrices-based output-only statistical damage localization and quantification on a frame structure," Sep. 2017.