Proceedings of the ASME 2011 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2011
August 12 – August 15, 2012, Chicago, IL, USA

IDETC 2012-71337

REPRESENTATION: STRUCTURAL COMPLEXITY OF ASSEMBLIES TO CREATE NEURAL NETWORK BASED ASSEMBLY TIME ESTIMATION MODELS

Michael G. Miller
Research Assistant
Department of Mechanical Engineering
Clemson University
Clemson, South Carolina 29634-0921
mm3@clemson.edu

James L. Mathieson
Research Assistant
Department of Mechanical Engineering
Clemson University
Clemson, South Carolina 29634-0921
jmathie@clemson.edu

Joshua D. Summers
Associate Professor
Department of Mechanical Engineering
Clemson University
Clemson, South Carolina 29634-0921
jsummers@clemson.edu

Gregory M. Mocko
Assistant Professor
Department of Mechanical Engineering
Clemson University
Clemson, South Carolina 29634-0921
gmocko@clemson.edu

ABSTRACT

Assembly time estimation is traditionally a time intensive manual process requiring detailed geometric and process information to be available to a human designer. As a result of these factors, assembly time estimation is rarely applied during early design iterations. This paper explores the possibility that the assembly time estimation process can be automated while reducing the level of design detail required. The approach presented here trains artificial neural networks (ANNs) to estimate the assembly times of vehicle sub-assemblies at various stages using properties of the connectivity graph at that point as input data. Effectiveness of estimation is evaluated based on the distribution of estimates provided by a population of ANNs trained on the same input data using varying initial conditions. Results suggest that the method presented here can complete the time estimation of an assembly process with +/- 15% error given an initial sample of manually estimated times for the given sub-assembly.

Keywords: graphs, DFA, artificial neural networks

1 ASSEMBLY TIME ESTIMATION

With the evolution of the Ford manufacturing system and the Toyota Production System in the 1950's, managers became more focused on eliminating waste and therefore paying attention to every work sequence and to every second that people, material, and machines were idle. This focus became the foundation for lean manufacturing, which resulted in an emphasis on identifying and eliminating waste, while increasing production and quality [1]. This trend however required the standardization of assembly time estimation procedures and has led to the development of many methods for predicting product assembly time [2,3,4].

Assembly time estimation is important to the field of design for assembly because it allows designers to predict, with varying degrees of accuracy, not only how long a product will take to assemble but also to compare assembly times between different design solutions. Increasingly stiff global competition requires manufacturing enterprises to compete in terms of quality, cost, and time to market of their products [5,6,7]. It is essential for these firms to understand the costs associated with the manufacture of their products in order to operate efficiently and competitively [7].

It is increasingly apparent that this understanding must be achieved earlier into the design phase. The possibility of influencing a product's cost substantially decreases early in the product's life cycle while the cost of modifications increases drastically as the project advances. Thus, decisions made early in a product's life cycle have a significant impact on the overall costs of the product, although the bulk of the expenses are incurred after the design phase has been completed as shown in Figure 1 [8]. Studies have shown that the design phase is often responsible for defining up to 80% of the costs a product incurs throughout its life cycle, but the design phase itself consumes less than 15% of the budget [8,2,9].

As a product moves from the design phase to production, more information about the product and the assembly process becomes available. In addition, the available information increases in certitude. It is expected that as more information becomes available, assembly time estimates should become more accurate. For example, little information is available during conceptual design, often only requirements and conceptual sketches. As the design process continues, more data is collected, including information regarding the way parts are connected to each other.

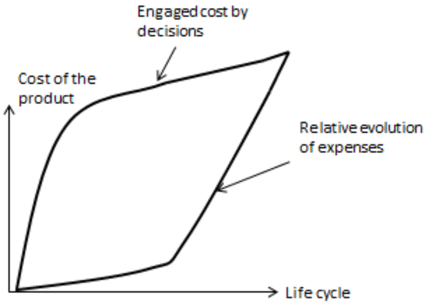

Figure 1: Cost engagements and expense occurrences throughout lifecycle [8]

Previous research has shown that assembly time estimates for consumer products can be predicted to within 16% of Boothroyd and Dewhurst assembly time estimates using only information about the part connectivity of the product [4]. This paper explores whether this approach can be adapted and applied to the automotive industry as illustrated in Figure 2.

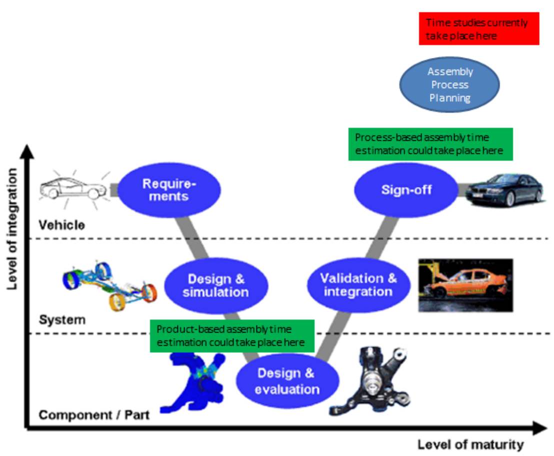

Figure 2: Automotive manufacturing product life-cycle, adapted from [10]

Assembly time estimates made at early stages of a product life-cycle are not expected to be as accurate as those which are conducted during production because there is substantially less information available to analyze. As a result, the estimates sought in this research are not intended to replace full scale assembly time studies which take place during production, but to supplement them by providing information earlier in the product life cycle.

1.1 Assembly Time Estimation Methods

In any effort to develop an assembly time estimation method, a means of assessing the assembly time of any given system must be employed. In the case of most traditional, late-stage assembly time estimation methods this takes the form of time studies requiring the direct observation of physical system assembly processes to a statistically significant replication count. However, for the development of an early-stage assembly time estimation method with an acceptable degree of inaccuracy, such exhaustive methods are both impractical and unnecessary. Rather, existing assembly time estimation methods can be used to acquire assembly time information which does not vary significantly enough from observed assembly time to skew results. In this paper, three assembly time estimation methods are discussed: methods-time measurement (MTM), Boothroyd and Dewhurst (B&D), and connectivity.

1.1.1 Methods-Time Measurement

Methods-time measurement, MTM, is a predetermined time study system in which operations are described by MTM "elements" [3]. The original MTM technique was time-consuming; however the method has since evolved to reduce the time required for analysis. It first became MTM-2 followed by MTM-3 and most recently MOST and MTM-UAS [11].

MTM analysis requires each motion of an operation to be analyzed. Analyzing an operation on this level allows for the user to easily identify obvious problems and non-value added motions [12]. The pre-determined times associated with specific motions have been determined in advance by statistical analysis. As an example, MTM-UAS defines seven basic motion take and place, place, helper equipment use, running (machine or equipment), motion cycles, body motions, and visual control [12].

An MTM style method is effective and relatively objective. However, the application of the method is labor intensive. The breakdown of an operation into its individual motions can take a significant amount of time. On the other hand, the objectivity and accuracy of the method lend MTM and MTM derivatives to application in industry. The majority of assembly time estimates used as targets in this research are the result of a custom MTM derivation developed by an automotive OEM.

1.1.2 Boothroyd and Dewhurst

The Boothroyd and Dewhurst method examines a part's handling and insertion difficulties when assigning an appropriate assembly time. A series of tables is used to determine an appropriate time for both handling and insertion of each part. The individual times within the tables have been developed from numerous time studies. The handling and insertion times selected from each set of tables are summed for each of the parts to determine the estimated assembly time for the product. This method has been proven to be beneficial, but is time consuming and includes subjective inputs [2,13]. Additionally, the information required for this analysis is often only present after the design is complete or near completion.

For example, to determine a handling time, the user must decide whether the part is easy or difficult to handle as well as the freedom of the part to rotate both parallel and perpendicular to its insertion point [14]. For insertion time, the user must determine whether holding down is required to maintain orientation, whether the part is easy or difficult to align, and whether or not resistance is present when inserting the part. Additionally, users must select whether the part and tool can easily access the desired location, as well as whether access is obstructed or vision is restricted.

The major drawbacks of both the tabular and software based Boothroyd and Dewhurst assembly time estimation methods include the need for significant amounts of information about both the product and the process and the number of subjective inputs required for the analysis.

1.1.3 Connectivity Assembly Time Estimation Method

The connectivity based assembly time estimation technique allows for the estimation of assembly time relying strictly on part connectivity information [4]. This method has only been tested on consumer appliances. However, it is likely that its use could be expanded to other industries including automotive manufacturing. For example, it may be possible to extend this method to an automotive sub-assembly and its part connectivity graph, such as those shown in Figure 3.

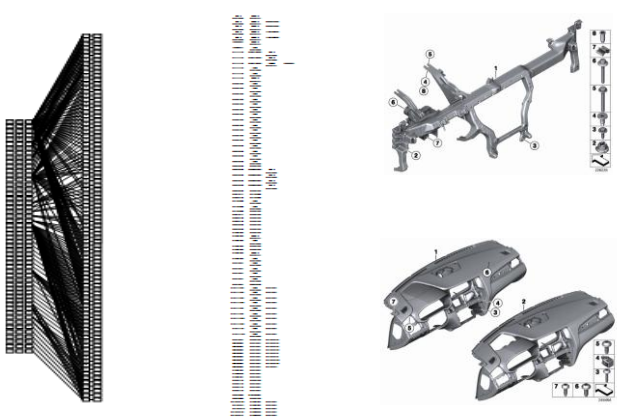

Figure 3: Bi-partite graph and tabular equivalent of automotive sub-assembly

The current connectivity method predicts assembly time based on a surrogate function on properties of the part connectivity graph he proposed model was developed using a manual pattern recognition approach, comparing and combining regression trends of average path length, number of elements, and path length density plotted against Boothroyd and Dewhurst assembly time estimates [4]. This regression based mapping was shown to be successful for consumer appliances as compared against Boothroyd and Dewhurst predicted assembly times. However, it has not been applied to products in different industries. The method could be fully objective with further definition of connections and has the potential to be fully automated.

1.1.4 Summary of Time Estimation Methods

With the exception of the connectivity method, the assembly time estimation methods which are currently available rely heavily on information that is not available until late in the design phase. This information includes required body movements, difficulty of handling and inserting parts, part size, part weight, part stickiness, and required connection order of parts. This poses a problem when users need information about assembly time earlier in the product's life cycle. For this reason, it is proposed that assembly time estimation could be obtained using part connectivity which is available earlier in the product life-cycle. One of the major motivations for this research is to use information more readily available early in a product's lifecycle to perform the analysis. Table 1 summarizes the information required in each of the assembly time estimation methods discussed so far. The information which can be obtained objectively has been shaded. This information can be obtained through automated and algorithmic software. The other, subjective information requires human interpretation and judgment.

Table 1: Questions asked of the designers for existing assembly time estimation techniques

MTM	Boothroyd and Dewhurst	Connectivity
What movements are necessary to perform assembly actions?	What order are parts connected in?	Which parts are connected to each other?
How difficult are parts to handle?	How difficult are parts to handle?	How many connection instances are present between parts?
How difficult are parts to insert?	How difficult are parts to insert?	
Part attributes such as envelope size, weight, stickiness, etc.	Part attributes such as envelope size, weight, stickiness, etc.	

The connectivity method has shown promise in estimating assembly times for consumer products based on time estimates derived from Boothroyd and Dewhurst time analysis. The research presented in this paper addresses the possibility of this method's refinement and application to the automotive manufacturing industry.

2 PRODUCT CONNECTIVITY INFORMATION

Information about the part connectivity of a product can be generated manually by reverse engineering a product or studying 3-D models and 2-D drawings [5,15,16]. Significant research has also been conducted on the ability to extract part connectivity information from computer-aided design models [17,18]. However, connectivity may be taken to mean different things.

One type of connectivity, the kind used for analysis in the connectivity method, has to do with what parts are in contact with other parts and the number of contact points [4]. Another form of part connection information is mating relationships defined in CAD models. These two types of information are related; however they are not the same thing. The focus of the research in this paper is connectivity in terms of physical part connections which may or may not have explicit mate conditions defined in the CAD environment. There are many ways to fully constrain parts in an assembly. Furthermore, different designers, or even the same designer, may constrain assemblies differently. For these reasons, physical part connections are considered in this research.

3 DATA COLLECTION

To begin the process of evaluating and refining the connectivity method, the connectivity graphs and associated MTM-based time estimates for a sample of systems are required. As discussed earlier, there are currently no automated means of generating connectivity information although current research in this field is ongoing [13]. As a result the connectivity graphs are obtained using a combination of observation of the assembly process, informal interviews with process associates, and information contained within CAD models.

Information collected for this analysis originates from the successive tasks performed in the assembly of one sub-assembly area for one specific vehicle. The assembly tasks are all within the instrument panel and take place between the attachment of the first part to a fixture and the completed sub-assembly's marriage to the vehicle.

One observation noted while collecting the information was that the graph did not change after the completion of every assembly task. This may be thought of as a limitation, but it could also be used advantageously. The limiting factor is obvious; the method is incapable of estimating an assembly time for tasks that do not result in any change in the connectivity of the parts. However, it is observed that the tasks that are not captured by the connectivity graphs are not truly value added activities. Although not explored in this paper, this finding could allow for automated connectivity analysis to identify non-value added activities. Some examples of these types of tasks are:

- Place cockpit sub-assembly broadcast sheet to the AGC with magnet.
- Remove two transport covers from upper flaps of the Heater/Aircon Low.
- Place in the recycle bin.

These three tasks do not affect the connectivity graph of the product. As a result, an assembly time cannot be estimated for these steps using an analysis of the connectivity graph.

The first process in the sub-assembly that results in a change to the connectivity graph is the attachment of the second part to the base part. The connectivity graph after the completion of this step is shown in Figure 4. This process involves attaching Part B to Part A and securing with two bolts (Bolt A and Bolt B) and two screws (Screw A and Screw B).

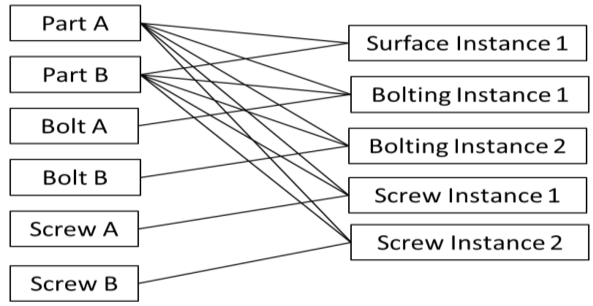

Figure 4: Connectivity graph after first set of assembly tasks

As the process is furthered, the connectivity graph continues to grow in size as more parts are introduced to the system and more connection instances are incorporated. The properties of the graph also change with the completion of more assembly processes. This trend is the focus of this research with the aim to map the properties of connectivity graphs to assembly times. The growth of the connectivity graph is shown in tabular form in Figure 5 and Figure 6.

IPC - 9201698	IHKA-9225727	
IPC - 9201698	IHKA-9225727	9906200A
IPC - 9201698	IHKA-9225727	9906200B
IPC - 9201698	IHKA-9225727	7142046A
IPC - 9201698	IHKA-9225727	7142046B

Figure 5: Tabular view of connectivity graph after one process

Figure 6 shows how the graph has increased in size after the completion of the first three assembly processes.

IPC - 9201698	IHKA-9225727		
IPC - 9201698	IHKA-9225727	9906200A	
IPC - 9201698	IHKA-9225727	9906200B	
IPC - 9201698	IHKA-9225727	7142046A	
IPC - 9201698	IHKA-9225727	7142046B	
7060601	IHKA-9225727		
7654321	IPC - 9201698		
9177178	9242146A	IPC - 9201698	
9177178	9242146B	IPC - 9201698	
9177178	9242146C	IHKA-9225727	7060601
9177178	IPC - 9201698		
9177178	IPC - 9201698		
9177178	IHKA-9225727		

Figure 6: Tabular view of connectivity graph after three processes

The data collection process continued for each of the non-option related tasks associated with the sub-assembly of the instrument panel. In total, 24 connectivity graphs and the associated assembly times for each activity were collected to be used in the analysis.

4 GRAPH PROPERTY ANALYSIS

The original connectivity method proposes that the properties of product connectivity graphs could be used to estimate assembly times for a given product [4]. The properties of a bi-partite graph, such as the one shown in Figure 3, are the basis for the analysis proposed in this paper as well as previous research [19]. The graph properties used in this research are based on the same as those proposed in the original assembly time estimation research based on connectivity [19]. These have been expanded to twenty nine properties, each falling into one of four main categories: size, interconnectivity, centrality, and decomposition to match the scheme developed in [19].

5 PERFORMANCE EVALUATION OF EXISTING CONNECTIVITY METHOD

The original connectivity method was designed and mapped to model assembly times for consumer products [4]. It was shown to map connectivity graphs to Boothroyd and Dewhurst assembly time estimates to within 16% [4]. Prior to establishing a new mapping scheme that could be applied to automotive industry assembly processes a study was conducted to determine the extent to which the original method could predict these times. The results of this study on twenty-four connectivity graphs shows an average error of twenty six percent with a range in errors between -134% and 352%. The plot of MTM-based assembly time estimates for these twenty-four graphs and the original connectivity estimates is shown in Figure 7. The function developed in the original model is not capable of accurately estimating assembly times in the automotive industry. Therefore, it is proposed that other, more complex techniques be explored to develop an appropriate mapping for this application.

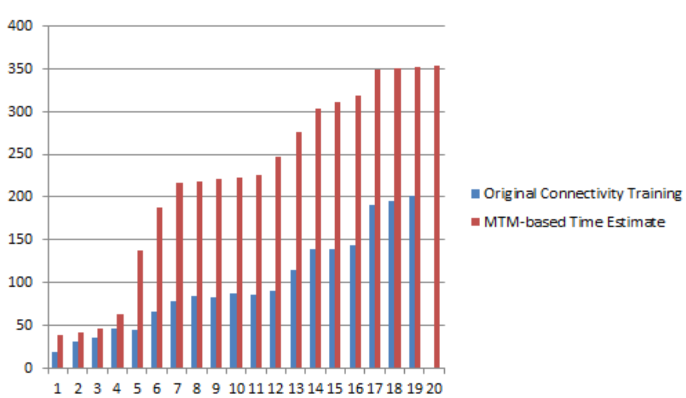

Figure 7: MTM-based estimates and original connectivity-based estimates

6 ARTIFICIAL NEURAL NETWORKS

The next step is the mapping of the connectivity graph properties to the MTM-based assembly time estimates. Artificial neural networks are chosen to explore this relationship due to their ability to perform nonlinear statistical modeling [20]. Other machine learning approaches, such as support vector machines and decision trees are ill-suited to this problem as they primarily perform a classification or clustering function and therefore do not provide for a continuous differentiable output. The advantages of neural networks include requiring less formal statistical training, the ability to detect complex nonlinear relationships between independent and dependent variables, the ability to discover all possible interactions between predictor variables, and the ability to use multiple training algorithms [20]. Artificial neural network analysis also has its disadvantages. These include its "black box" nature, the greater computational expense, the tendency to "over-fit", and the empirical nature of the development of the model [20,21]. As the purpose of this research is to develop a model which reliably generalizes between the input graph properties and the assembly time without a need for physical meaning between them, the "black-box" nature is acceptable. The issue of over-fitting is addressed in training by instituting an early stopping algorithm as well as withholding samples from training entirely to test generalization on non-training data.. Table 2 summarizes the results of a few studies about the applicability and performance of artificial neural networks as compared to other analysis methods. For a more comprehensive assessment of the literature see [21]

Table 2: Artificial neural network comparisons in literature

Application of ANN	What ANN was compared to	Conclusions	Ref.
Predict dynamic nonlinear systems	Statistical Models	ANNs provide satisfactory performance in forecasting	[21]
Forecasting	Box-Jenkins automatic forecasting expert system	Similar Results	[22]
Nonlinear statistical modeling of medical outcomes	Logistic regression	Neural Networks preferred when primary goal is outcome prediction	[20]
Cost Estimation of Steel Pipe Bending	Linear regression	Neural Network	[23]
Prediction of Commodity Prices	Logistic regression	ANNs are consistently better and find more turning points	[24]

Two critical factors for this analysis are the inputs and the targets. The input for the analysis is the vector of graph properties for each connectivity graph. It should be noted that the connectivity graph at any time represents all of the connections made up to that point in time. This includes the execution of all assembly tasks that are required to make all of the connections present in the connectivity graph. Therefore, the graph property vector for a connectivity graph is to be mapped to the total assembly time up to that point. To determine the time for an isolated assembly step or steps the estimated assembly time prior to that step must be subtracted from the total estimated time including the step.

The target for the mapping is the MTM-based assembly time estimate provided by an automotive OEM. These estimates are the result of a formal study performed by time study personnel. The studies are conducted using a company specific adaptation of MTM-UAS. Again, the goal of the methods proposed in this paper is not to replace the formal, late-stage time study. The aim of this research is to provide an assembly time estimate much earlier in the product life cycle and to enable the automation of such a method. Formal time studies come late in the product life-cycle, after production has begun. The added level of detail and analysis time results in a more accurate estimate. Therefore, the assembly time estimates provided by the OEM implementation of MTM-UAS are used as target values in this study.

The process of building the model scheme is shown in a more detailed manner in Figure 8. The graph property vectors representing 19 connectivity graphs and their associated assembly times are used as inputs and targets. Graph property vectors for 5 connectivity graphs and their associated assembly times are withheld from training for later validation. One hundred simulations are then performed for each of 189 different ANN architectures. Next, a probability density function is generated for each of the architectures. The architectures are then evaluated based on the probability of predicting assembly times to within fifteen percent of the target time. This is calculated by integrating the area under the probability density plot between the upper and lower fifteen percent bounds. Finally, a combination of the 100 predicted assembly times from the five best performing neural networks is used to generate a probability density plot. It is expected that the combination of the top performing architectures will enable the model to more accurately predict assembly times for different vehicle areas.

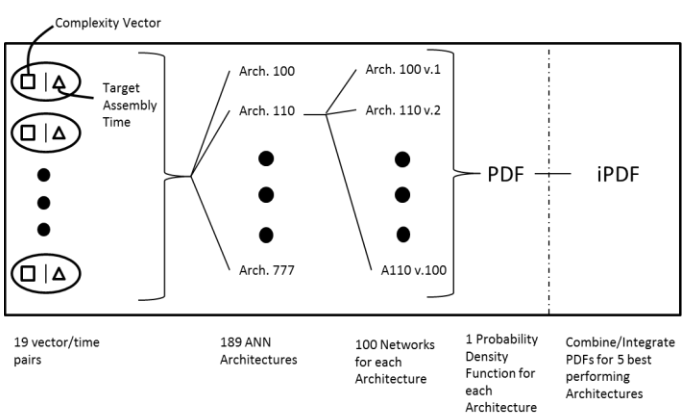

Figure 8: Model building process

Once the model is built and tested in order to understand the degree of accuracy, it can be used to estimate assembly times. The process of using the model is illustrated in Figure 9. This figure shows that a graph property vector representing a connectivity graph is used as the input to the model. Then, the model is simulated using the five top-performing ANN architectures. Finally, a probability density function is generated using 500 ANN replications (100 from each of the five architectures). This probability density function can be used to gain an understanding of the predicted assembly time.

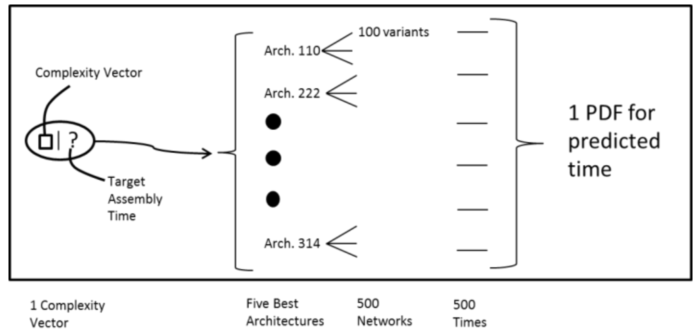

Figure 9: Process of using the model

Roughly 20%, or 5 of the 24 graphs, are omitted from training for external testing. This leaves 19 of the 24 data points to be used as inputs and targets for the analysis. The remaining five are used after the appropriate ANN architectures are selected to determine the accuracy to which the network was capable of mapping connectivity graphs to MTM-based assembly time estimates.

One hundred and eighty nine different ANN architectures were simulated in order to identify the most appropriate for this mapping. These architectures range from a single layer with a single neuron to three layers with five neurons in each layer. Architectures with one layer were simulated with a neuron count ranging from one to fifteen. Architectures with two layers were simulated with one neuron each up to seven neurons in each layer. Finally, architectures with three layers were simulated with combinations of up to five neurons. This caps the number of hidden layer units at 15. This places the hidden unit count between the input unit count and output unit count, thus promoting generalization of the output [25].

An ANN will not generate the exact same mapping even when given the same inputs and outputs due to different initial conditions in each training. The most notable of these variations is the use of an early-stopping validation algorithm which withholds a 15% subset of the training data for use in testing that while training progresses generalization also continues to improve. Without this measure, the network could be expected to consistently over-fit to the training data. As the subset of the training data to be used for early-stopping is selected at random, the training of each architecture is executed for 100 replications of the ANN. This allows for the generation of probability density functions describing the typical behavior of the 189 ANN architectures. A probability density plot for an ANN architecture consisting of three layers with five neurons in the first two layers and one neuron in the third layer evaluated for one connectivity graph is shown in Figure 10. An equivalent plot is generated for all combinations of ANN architectures and connectivity graphs.

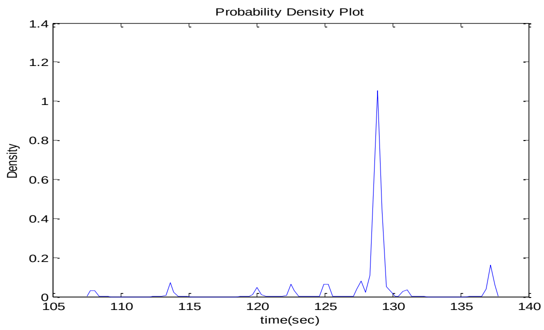

Figure 10: Probability density plot for connectivity graph #5 and ANN structure 134

The next step is to analyze where the probability function lies in relation to the target value and acceptable range of assembly times. Figure 11 shows the probability density plot as well as the target value and associated 15% range of the target values for the assembly. The mean predicted value is shown by the red dotted line while the black dotted line shows the target assembly time value. It is noted that the predicted values fall well within the 15% acceptable range.

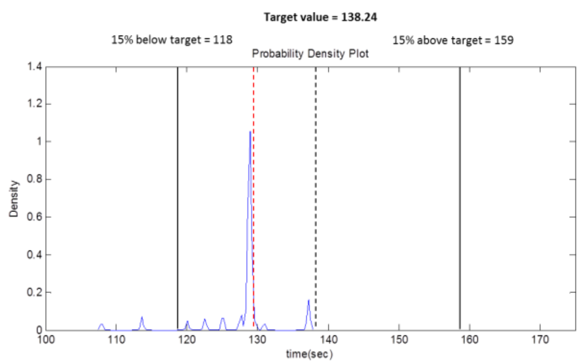

Figure 11: Probability density plot with target value and 15% range

Next, the artificial neural network architecture with the highest probability of estimation to within fifteen percent of the target is identified to be used for future assembly time estimation. This probability is based on the assembly time estimation for the five connectivity graphs which are omitted from training. It should be noted that four of the ANN architectures resulted in probabilities of greater than one, the result of integration errors in computation. To determine the probability, the area under the probability density plot between the upper and lower limits was calculated using trapezoidal approximate integration. It should also be noted that some of the architectures result in a probability of zero. This suggests that none of the simulated assembly times fall within a fifteen percent range of the target times.

The most appropriate ANN architecture was determined based on the data presented in Table 3. Columns two and three represent the probability that the minimum and mean of the estimation will be within 15% of the target value for the five validation sets. The first column identifies the ANN architecture. It can be seen that each of these have a high probability of estimating the assembly time to within 15% of the MTM-based assembly time estimate. Since there is not a significant difference for any of the cases between performance on the predicted mean and minimum, the ANN architectures with the highest mean probability are selected. The top ANN architecture consists of three layers with three neurons in the first layer, four in the second, and five neurons in the third layer.

Table 3: Evaluation of ANN structure performance

ANN Architecture	Minimum Probability of Estimation to within 15% of Target	Mean Probability of Estimation to within 15% of Target
134-[3,4,5]	0.99991	0.99994
188-[5,5,4]	0.99990	0.99992
153-[4,3,4]	0.99983	0.99990
157-[4,4,3]	0.99983	0.99990
77-[1,3,3]	0.99985	0.99989
32-[3,3]	0.99977	0.99987
170-[5,2,1]	0.99966	0.99986
69-[1,1,5]	0.99970	0.99983

7 RESULTS

As previously mentioned, 5 of the 24 data sets are omitted from training for validation separate from any validation set generated by the training algorithm. These data points are used to test the ability of the selected ANN architecture to generalize new data. The top network is trained two additional times using the selected architecture while omitting different sets of five data points. During the original training and testing, the five largest connectivity graphs are omitted to use later as test points. This test is used to determine the forward prediction of the model. The second training omits every fifth data point (when ranked in terms of connectivity graph size) beginning with the smallest connectivity graph. The third validation set consists of every fifth connectivity graph starting with the fifth. The second and third validation sets are used to determine the applicability of the model to a wide range of graph sizes after training on a representative sample. The final validation set consists of using a combination of the top performing architectures.

7.1 First Validation Set (Last Five Omitted)

The ANN population is trained using the first 19 data points as inputs and targets and the top ANN architecture. This network was then simulated with the input being the graph property vectors of the final five data points. The results of the ANN validation simulation are shown in Table 4. This table shows the estimated time from the MTM-based time study provided by the automotive OEM and the estimated time by the

ANN model for each data point. The estimate from the model is the mean output from a population of 100 ANN replications. The error represents the percentage error of the model as compared to the formal time study estimate. The final three columns show the probability that each of the 100 ANN replications predict a time to within a specified percentage of the formal time study estimate. The results in Table 4 show that this model is capable of forward prediction within an area of the vehicle. In other words, the ANN trained on a set of smaller connectivity graphs is capable of predicting assembly times for graphs larger than those used in the training.

Table 4: Prediction results for first validation set

Graph	MTM [s]	ANN [s]	Error	Probability of prediction within:		
				10%	5%	1%
20	352.92	350.30	-0.7%	1	1	.52
21	362.64	370.18	2.1%	1	.99	.27
22	366.96	384.45	4.8%	1	.53	0
23	376.62	386.09	2.5%	1	.72	.18
24	392.94	386.94	-1.5%	.99	.58	.12

The results of this validation set show a successful mapping between the complexity of connectivity graphs and MTM-based assembly time estimates for a specialized case. This validation is only applicable to forward prediction of assembly times within a specific vehicle area. This validation does not imply anything about mapping of assembly times in other parts of a vehicle. In addition it does not speak to the model's ability to predict assembly times for connectivity graphs which are smaller than the input graphs or for a wide range of sizes of connectivity graphs. For this reason, a second validation is performed and discussed in Section 7.2.

7.2 Second Validation Set (Every Fifth Omitted Starting with Graph 1)

The second validation set seeks to explore the model's ability to predict assembly times for a wider range of connectivity graph sizes for cases in which the model is trained on a representative sample of the population. The ANN population for this case was trained using 19 of the 24 collected data points as inputs and targets and top performing ANN architecture. This ANN population was then simulated with the input being the graph property vectors of every fifth data point, or connectivity graph, starting with the smallest.

The results of the second neural network validation simulation are shown in Table 5 using the format established in Table 4. As shown in the table, the only connectivity graph which was not predicted with an error of less than 15% was the first. This connectivity graph is the smallest graph in any of the collected data, and is consequently smaller than the training set. This error, along with the low probability of estimation to within 10% highlights this procedure's lack of ability to predict assembly times for graphs which are smaller than those used in the training.

Table 5: Prediction results for second validation set

Graph	MTM [s]	ANN [s]	Error	Probability of prediction within:		
				10%	5%	1%
1	39.3	45.73	16.4%	.23	.11	.03
6	187.38	202.29	8.0%	1	0	0
11	226.26	225.10	-0.5%	1	.96	.02
16	318.78	315.40	-1.1%	1	1	.47
21	362.64	372.90	2.8%	1	1	0

The results of the second validation set show great potential for the deployment of the method. However, it also identifies a weakness in the model's ability to predict assembly times for graphs which are smaller than those present in the network's training. As a result of this observation, the goal of the third validation set is to determine the model's applicability when trained on a representative sample of the population with upper and lower bounded data.

7.3 Third Validation Set (Every Fifth Omitted Starting With Graph 5)

The third validation set omits every fifth graph in terms of size, but starts with graph 5. This validation seeks to examine the capabilities of the method when trained on a representative sample which contains the upper and lower bounded graph sizes. In this validation, the graphs omitted for validation are between the largest and smallest graphs used in the training of the ANN population.

The results of the third ANN validation set are shown in Table 6, in a format similar to that established in the previous validation sets. As shown in the table, 100% of the predictions for this validation are within 25% of the target value. Similarly, for every graph except graph 5, all of the simulations are within 15%.

Table 6: Prediction results for third validation set

Graph	MTM [s]	ANN [s]	Error	Probability of prediction to within:				
				25%	15%	10%	5%	1%
5	138.24	122.67	-11.3%	1	.67	.44	.22	.04
10	223.20	228.28	2.3%	1	1	1	.95	.29
15	311.04	311.43	0.1%	1	1	1	1	.98
20	352.92	351.65	-0.4%	1	1	1	1	1
24	392.94	408.98	4.1%	1	1	1	.66	0

The third and final validation set demonstrates that an ANN population can successfully predict assembly times when trained on a representative sample of the population and when the extreme data points in terms of graph size are included in the training.

7.4 Fourth Validation Set (Every Fifth Held Back Starting with Graph 5) using best five architectures

In the final validation set within the instrument panel area of the vehicle a combination of the top five performing artificial neural networks (134, 188, 153, 157, and 77) is used. To obtain the results for this analysis, an ANN population with equal members of each of the five architectures is simulated for the test cases. For each architecture 100 replications exist within the population, resulting in 500 predicted times for each test case. Finally, a probability density plot is generated using the 500 assembly times. Table 7 shows a summary of the results of the fourth validation set.

The results of the fourth validation set are very accurate. The mean predicted values for each of the test points are within 2.5 percent of the target assembly time. This suggests that using a combination of top performing architectures is helpful in successfully mapping the complexity of connectivity graphs to assembly times.

Table 7: Prediction results for fourth validation set

Graph	MTM [s]	ANN [s]	% Error	Probability of prediction to within:		
				10%	5%	1%
5	138.24	138.15	-0.1%	.51	.08	.02
10	223.20	223.20	0.9%	1	1	.37
15	311.04	311.04	1.1%	1	1	.46
20	352.92	352.92	-0.3%	1	.94	.94
24	392.94	392.94	-2.2%	1	.81	.01

Table 8 shows a summary of the validation results. It is shown that the results within a particular vehicle area are accurate. However, the best results are obtained when the five top-performing neural networks are used to simulate the assembly times. This can be seen in the last column where the maximum error is less than 2.5 percent.

Table 8: Summary of validation results

Test	Last 5 graphs	Every 5th graph, starting with #1	Every 5th graph, starting with #5	Every 5th graph, best 5 architectures
1	-0.7%	16.4%	-11.3%	-0.1%
2	2.1%	8.0%	2.3%	0.9%
3	4.8%	-0.5%	0.1%	1.1%
4	2.5%	-1.1%	-0.4%	-0.3%
5	-1.5%	2.8%	4.1%	-2.2%

8 EXTERNAL GENERALIZATION

The results in Section 7 show that the connectivity method is capable of predicting assembly times when tested on the vehicle area used for the ANN training. However, the application of this mapping to other areas of the vehicle or to non-automotive assemblies has not yet been explored.

8.1 Application to other parts of the vehicle

The first question to be addressed in external generalization is whether the ANN trained only on the instrument panel is capable of predicting assembly times for other parts of the vehicle. The insulating panel will serve as the other vehicle area for this analysis.

The ANN developed in Section 7 was simulated to predict assembly times for the graph property vectors of the insulating panel's connectivity graphs. The results from this analysis are shown in Table 9. The errors in this validation set range from negative forty-two percent to 435%. This lack of consistency suggests that it is difficult to use a model trained on one specific vehicle area to estimate assembly times for a different vehicle area.

Table 9: External Generalization Results

Graph	MTM [s]	ANN [s]	% Error
1	20	69	231
2	27	15	-42
3	30	34	15
4	34	184	435
5	47	181	284
6	60	201	235
7	80	259	223
8	95	301	214

The results shown in Table 9 yield the second question which addresses whether a neural network trained on complexity vectors and times from multiple vehicle areas is capable of producing accurate estimates for the different vehicle areas. To address this question an ANN population 7 is trained on the instrument panel as performed in Section 7 as well as three of the eight connectivity graphs collected from the insulating panel. Table 10 shows the results of this analysis.

Table 10: Connectivity results for insulating panel assembly

Graph	MTM [s]	ANN [s]	% Error
2	27.18	47.67	75%
4	34.56	97.97	183%
5	47.34	84.52	79%
7	80.28	54.66	-32%
8	95.94	67.12	-30%

The results show that estimation of assembly times for multiple vehicle areas is possible with some degree of accuracy. However, the results are not as accurate as those within the same vehicle area as the training set. This lower level of accuracy is likely due to the fact that a much larger number of instrument panel processes was used than insulating panel processes.

Since the training set now includes data from the insulating panel, the accuracy of the estimation of instrument panel assembly times may have been reduced. The assembly time estimation for the instrument panel processes both before and after the addition of insulating panel processes are presented in Table 11. This shows that when the ANN population is trained on insulating and instrument panel assemblies the accuracy of the assembly time estimates for the instrument panel is not decreased. This may suggest that a higher number of processes used in the neural network results in a higher degree of accuracy in assembly time prediction regardless of the vehicle areas analyzed.

Table 11: Connectivity results for instrument panel assembly

Graph	ANN [s]	MTM [s]	% Error	% Error (IP-only Training)
IPA-5	199	187	6%	12%
IPA-10	244	226	8%	-3%
IPA-15	313	318	-2%	-3%
IPA-20	361	362	0%	-2%

8.2 Application to Consumer Products

The next step is to determine the applicability of the newly developed ANN model to assembly time estimation outside of automotive assembly processes. The original connectivity training was developed for use on consumer product assemblies [19]. To determine the new model's ability to predict assembly times for consumer products, the model is tested on three products used in the initial connectivity research including a mixer, a chopper, and a Tweel™ prototype.

The results of this analysis are presented in Table 12. This shows that it is necessary to train the model on data specific to the application it is to be used for. Furthermore, it is likely that different automotive OEMs would need to train the model specifically for application in the respective company.

Table 12: Results of model application to non-automotive assemblies

Graph	B&D [s]	ANN [s]	% Error
Mixer	136	180	32%
Tweel™	13561	228	-98%
Chopper	228	35	-84%

9 CONCLUSIONS & FUTURE WORK

This paper has explored the possibility that an automatable, early-stage assembly time estimation model based on part connectivity can be developed for the automotive industry. It has been found that existing regression based assembly time estimation methods using connectivity information are not extendable to automotive assemblies. To address this, an artificial neural network mapping approach incorporating populations of ANNs is proposed and evaluated. This evaluation has shown that the ANN approach is well suited to predicting assembly times for intermediate steps when trained on a representative sample containing the upper and lower bounds, suggesting its possible use as an accelerating supplement for formal time studies. Fair performance is also observed for predicting the assembly time of steps beyond the upper bound, indicative of progress towards the goal of an automated early-stage assembly time estimation tool.

It is recommended that a model to be used for assembly time estimation be trained on a set of graph property vectors representing the upper and lower bounded connectivity graphs in addition to a representative sample of intermediate graphs. Additionally, a population consisting of five or more of the top performing network architectures should be used to generate a probability density plot representing the estimated assembly time. It is shown that predictions generated from this model are only applicable to vehicle areas on which it has been trained and is not viable for direct use in other industries. Further, it is hypothesized that the model would also be company-specific and fail to generalize across automotive OEMs.

The results of this study suggest that a higher number of training data points representing a sample of each of the vehicle areas may result in a model which is capable of accurately predicting assembly times for all areas of the vehicle. The development of such a model will require significantly more effort due to the large number of samples required. A definite rule does not exist for sample sizes, but the size of the training set depends on the network structure, training method, and the complexity of the problem [26]. However, it is possible that the result would be a product-based assembly time estimation model capable of providing accurate results early in the automotive product life-cycle.

AKNOWLEDGEMENTS

This work was supported with funds from BMW (2008367). The views presented here do not necessarily represent those of the funding sponsors.

REFERENCES

[1] Mary S. Spann, Mel Adams, Maruf Rahman, Hank Czarnecki, and Bernard J. Schroer, "Transferring Lean Manufacturing to Small Manufacturers: The Role of NIST-MEP," in *United States Association for Small Business and Entrepreneurship Conference*, San Diego, 1999.

[2] Geoffrey Boothroyd, Peter Dewhurst, and Winston A. Knight, *Product Design for Manufacture and Assembly*. Boca Raton: CRC Press, 2011.

[3] C.K. Choi and W.H. Ip, "A Comparison of MTM and RTM," *Work Study*, vol. 48, no. 2, pp. 57-61, 1999.

[4] James L. Mathieson, Bradley A. Wallace, and Joshua D. Summers, "Assembly Time Modeling Through Connective Complexity Metrics," in *International Conference on Manufacturing Automation*, Hong Kong, 2010.

[5] Kevin N. Otto and Kristin L. Wood, "Product Evolution: A Reverse Engineering and Redesign Methodology," *Research in Engineering Design*, vol. 10, pp. 226-243, 1998.

[6] Prabhu Shankar, *Development of a Design Method to Reduce Change Propagation Effects*. Clemson, 2011.

[7] David Ben-Arieh and Li Qian, "Activity-based Cost Management for Design and Development Stage," *International Journal of Production Economics*, no. 83, pp. 169-183, 2003.

[8] P. Duverlie and J.M. Castelain, "Cost Estimation During Design Step: Parametric Method Versus Case based Reasoning Method," *International Journal of Advanced*

Manufacturing Technology, no. 15, pp. 895-906, 1999.

[9] Jianxin Jiao and Mitchell M. Tseng, "A Pragmatic Approach to Product Costing Based on Standard Time Estimation," *International Journal of Operations and Production Management*, vol. 19, no. 7, pp. 738-755, 1999.

[10] Julian Weber, *Automotive Development Processes: Processes for Successful Customer Oriented Vehicle Development*.: Springer, 2009.

[11] J, Laring, M. Forsman, R. Kadefors, and R Ortengren, "MTM-based Ergonomic Workload Analysis," *International Journal of Industrial Ergonomics*, vol. 30, pp. 135-148, 2002.

[12] Mehmet Cakmakci and Mahmut Kemal Karasu, "Set-up Time Reduction Process and Integrated Predetermined Time System MTM-UAS: A Study of Application in a Large Size Company of Automobile Industry," *International Journal of Manufacturing Technology*, vol. 33, pp. 334-344, 2007.

[13] Eric Ownesby, Aravind Shanthakumar, Vikrant Rayate, Essam Namouz, and Joshua D. Summers, "Evaluation and Comparison of Two Design for Assembly Methods: Subjectivity of Information Inputs," in *International Design and Technical Conferences*, Washington D.C., 2011.

[14] Joshua D. Summers. (2010) Clemson Engineering Design Applications and Research. [Online]. http://www.clemson.edu/ces/cedar/images/4/45/2009-01-ME455-Lecture-02-DFA.pdf

[15] Mark Roy Snider, Extended Toolset for Reverse Engineering to Support Lightweight Engineering, 2006.

[16] M Snider, J. Summers, S. Teegavarapu, and Mocko G., "Database Support for Reverse Engineering, Product Teardown, and Redesign as Integrated into a Mechanical Engineering Course," *ASEE Computers in Education Journal*, 2007.

[17] Arun Mathew and C.S.P. Rao, "A CAD System for Extraction of Mating Features in an Assembly," *Assembly Automation*, vol. 30, no. 2, pp. 142-146, 2010.

[18] Kedar Sambhoos, Bahattin Koc, and Rakesh Nagi, "Extracting Assembly Mating Graphs for Assembly Variant Design," *Journal of Computing and Information Science in Engineering*, vol. 9, pp. 034501-1-034501-9, 2009.

[19] James L. Mathieson, Connective Complexity Methods for Analysis and Prediction in Engineering Design, May 2011.

[20] Jack V. Tu, "Advantages and Disadvantages of Using Artificial Neural Networks versus Logistic Regression for Predicting Medical Outcomes," *Journal of Clinical Epidemiol*, vol. 49, no. 11, pp. 1225-1231, 1996.

[21] Guoqiang Zhang, B. Eddy Patuwo, and Michael Y. Hu, "Forecasting with Artificial Neural Network," *International Journal of Forecasting*, vol. 14, pp. 35-62, 1998.

[22] Ramesh Sharda and Rejendra B. Patil, "Connectionist Approach to Time Series Prediction: An Empirical Test," *Journal of Intelligent Manufacturing*, vol. 3, pp. 317-323, 1992.

[23] Avraham Shtub and Ronen Versano, "Estimating the Cost of Steel Pipe Bending, a Comparison Between Neural Networks and Regression Analysis," *International Journal of Production Economics*, vol. 62, pp. 201-207, 1999.

[24] Nowrouz Kohzadi, Milton S. Boyd, Bahman Kermanshahi, and Iebeling Kaastra, "A Comparison of Artificial Neural Network and Time Series Models for Forecasting Commodity Prices," *Neurocomputing*, vol. 10, pp. 169-181, 1996.

[25] A. Blum, *Neural networks in C++: an object-oriented framework for building connectionist systems, Volume 1*.: Wiley, 1992.

[26] Guoqiang Zhang, B. Eddy Patuwo, and Michael Y. Hu, "Forecasting with artificial neural networks," *International Journal of Forecasting*, no. 14, pp. 35-62, July 1998.

