Leverage Artificial Intelligence to Learn, Optimize, and Win (LAILOW) for the Marine Maintenance and Supply Complex System

Ying Zhao
Naval Postgraduate School
Monterey, CA 93943, USA
yzhao@nps.edu

Gabe E. Mata
MARCORSYSCOM
Albany, GA

***Abstract*—A complex enterprise includes multiple subsystems and organizations. The U.S. Marine Corps (USMC) maintenance and supply chain is a complex enterprise and exemplifies a socio-technological infrastructures. It is imperative for the USMC to adopt more advanced data sciences including ML/AI techniques to the entire spectrum or end-to-end (E2E) logistic planning as a complex enterprise including maintenance, supply, transportation, health services, general engineering, and finance. In this paper, we first review an overall framework of leveraging artificial Intelligence to learn, optimize, and win (LAILOW) for a complex enterprise, and then show how a LAILOW framework is applied to the USMC maintenance and supply chain data as a use case. We also compare various machine learning (ML) algorithms such as supervised machine learning/predictive models and unsupervised machine learning algorithms such as lexical link analysis (LLA). The contribution of the paper is that LLA computes stable and sensitive components of a complex system with respective to a perturbation. LLA allows to discover and search for associations, predict probability of demand and fail rates, prepare spare parts, and improve operational availability and readiness.**

***Index Terms*—supervised machine learning, predictive models, unsupervised machine learning, lexical link analysis, LLA, association patterns, data mining, maintenance and supply chain, complex system, leverage artificial intelligence to learn, optimize, and win, LAILOW**

I. Introduction

A complex enterprise with multiple subsystems and organizations is omnipresent. A Navy fleet is a complex enterprise and the USMC maintenance and supply system is a complex enterprise as well. A complex enterprise contains a myriad of business processes as subsystems that can be either sequential or parallel. A complex enterprise needs trusted AI to achieve automation, foster collaborations, and win competitions. We can leverage data sciences and advanced ML/AI techniques for complex enterprises.

IEEE/ACM ASONAM 2020, December 7-10, 2020

According to [1], [2], uncertainty, unknown operation conditions are the key challenges for the U.S. Marine Corps (USMC) maintenance and supply chain. For example, the uncertainty of the reliability of assets (predicted failure/remaining useful life (RUL)) has caused the USMC to maintain and operate with excess equipment and supplies.

Deep data analytics including machine learning (ML) have been used by the USMC to address variety of challenges. For example, predictive methods of supervised ML methods have been used to predict equipment reliability and probability of failure, therefore infer numbers of spare parts to improve stock performance and synchronize budget execution.

It is imperative for USMC to adopt more advanced data sciences including ML/AI techniques to focus on the entire spectrum or end-to-end (E2E) logistic planning for the complex enterprise of maintenance, supply, transportation, health services, general engineering, and finance. The ultimate goal is to enhance the total force readiness and project combat power across the whole range of military operations and spectrum of conflict at any time.

In this paper, we first review an overall framework of leveraging artificial Intelligence to learn, optimize, and win (LAILOW) for a complex enterprise, and then show how the LAILOW framework is applied to the USMC maintenance and supply system as a user case. The contribution of this paper is that we show an unsupervised machine learning method lexical link analysis (LLA) is powerful to discover the associations and networks of items for a logistics and supply system. We then show that the discovered associations combined with the predictive models in a game-theoretic set up of LAILOW propagates and reaches an equilibrium of total item demand prediction for the whole system when a demand perturbation to the system is introduced by high-impact and low-occurrence items. The methodology can be extended to other complex systems to predict stable states of a whole system with subsystem interactions and dynamics.

II. LAILOW Framework

A LAILOW framework can be summarized as follows:

- Step 1 - Learn: LAILOW first learns patterns and rules from historical data using data mining, big data, and machine learning techniques. Patterns and rules describe the correlations, associations, predictive patterns, or transition patterns among business processes and subsystems. For example, A USMC's unit structure includes a table of organization and equipment and a large number of items and parts to support the core equipment. Each core equipment and parts has specific need for the (X) duration and frequency of manpower and equipment for maintenance. One should first perform data mining, exploratory analysis, visualization, and causal learning models to understand the time and processes needed for maintenance services and pinpoint the causes for high cost or slow processing areas such as failure rates, demand patterns, or available manpower. The resulted machine learning models can be then used for predicting the desired effects for the future data.

  A wide range of tools are needed to data-mine the historical data to address the core attributes, for example, the probability of fail (POF) of a part or equipment. An unsupervised machine learning method called Lexical Link Analysis (LLA) is used in this paper to discover patterns of parts and items that are demanded together which is related to POF. The patterns are developed into networks and graphs for parts. The networks and graphs can be then combined with predictive models to provide a global, holistic, and associated view of spare parts needed should one or multiple new conditions occur as perturbations to the complex enterprise.
- Step 2 – Optimize: Based on the predictive patterns and association patterns developed from Step 1, LAILOW optimizes the measures of effectiveness (MOEs) or the measures of performances (MOPs), defined by business decision makers, by searching through better possible courses of actions for future requirements.

  For example, optimization methods have been used to optimize and maximize the throughput capacity at each node of the USMC maintenance and supply system to match the combat power and readiness necessary. A logistician/planner would have to assess the physical network analysis (PNA) or also known as logistical network analysis, to determine and predict the rate of flow in various new operation conditions. Thus, planners would be able to forecast better the rate of combat power entering into an area of operation to avoid congestion or delays in operations. Optimizing the stocking parts based on the ML predictive maintenance models helps prevent ordering surplus or unnecessary components for "just in case" or better known as the Iron Mountain Concept [1], [2].
- Step 3 – Win: LAILOW represents a complex enterprise with a logistician/planner as a self-player in a game environment in real-time and suggests winning actions based on the nature of an opponent. The opponent can be new environmental and/or operational conditions. This requires war game type of ML/AI and simulation tools to simulate the response in new environmental/operational conditions.

  This is also related to the Analysis of Alternatives (AoA), conducted as simulations and what if analysis to create added conditions for new USMC operational conditions such as an IED blast, desert environment, and corrosion considerations. In order to compute "delta" for replenish based on the perturbations to the complex system, gathered intelligence data, knowledge and rules from SMEs, and existing predictive and engineering models can feed to the system to provide extra data to predict such "delta." It is important to perform this type of AoA or simulations or games since there might be no historical data available for new conditions, traditional predictive modeling analysis might not be directly applicable. So LAILOW can create such new conditions as an opponent to the self-player of a war game to see how the predictive models combined with association patterns of spare parts can adjust better predictions for the new conditions.

  Game theory and generative adversarial networks (GAN) [3], [4] have been interestingly considered in many commercial ML/AI applications such as generating simulation and synthetic data that are not easily observed in real life to improve the performances of ML/AI systems. Combined with simulation and war game tools as in the LAILOW framework, it is possible to address, learn, and self-practice the situations that are never seen before such as uncertain, no data, and adversarial conditions for a complex enterprise.

III. Uncertainty, Perturbation, Association, and Cascade Effects

The probability of fail of a part can be affected by many factors. We need to consider the uncertainty, disruption and perturbation that can impact the logistics plans as a whole. For example, uncertainty factors related to environment and events in wide geographic areas, such as, weather change or mission change from a path to another, or a sudden event can cause a perturbation, disruption, and cascade effects for the whole system.

The probability of fail is also embedded in a long chain of historical maintenance and supply data. A failed part can be fixed without ordering a new part. A part order frequency or probability of demand (POD) in the historical supply data reflects the demand if a part can not be repaired within a certain period of time. The probability of demand (POD) in the supply data can only reflect partial probability of fail (POF). Sometimes, the business practice may also result in patterns of POD that can not be explainable using POF, for example, some USMC units may tend to order particular parts more than other units.

The complexity of predicting total probability of fail for a large list of the items calls for the integration of methods in data fusion, data mining, optimization, and game theory when facing particular uncertainty and perturbation. In this paper, we focus on the maintenance and demand/supply processes as the relevance to the spare parts with an initial data integration.

IV. Discovering Item Association Network Using Lexical Link Analysis (LLA)

A key contribution of this paper is to apply Lexical Link Analysis [8] for predicting probability of demand (POD). LLA is an unsupervised machine learning method and describes the characteristics of a complex system using a list of attributes or features, or specific vocabularies or lexical terms. Because the potentially vast number of lexical terms from big data, the model can be viewed as a deep model for big data. For example, we can describe a system using word pairs or bi-grams as lexical terms extracted from text data. LLA automatically discovers word pairs, and displays them as word pair networks.

Bi-grams allow LLA to be extended to numerical or categorical data. For example, using structured data, such as attributes from the USMC maintenance and supply databases, we discretize numeric attributes and categorize their values to word-like features. The word pair model can further be extended to a context-concept-cluster model [9]. A context can represent a location, a time point, or an object shared across data sources.

We use LLA for the structured data of the USMC maintenance and supply databases to discover associations among items/parts in demand, therefore infer and predict the spare parts needed to improve the total readiness level.

Fig. 2 shows conceptually how the associations and correlations are discovered by LLA. We anticipate the demand change (DC) for an item/part i might come from two types of sources: Type 1): A collection of outside perturbations such as the change of missions or new operational conditions; and Type 2): Item associations with other items where the associations could be due to physical linkages or linked demand based on past business practices. If an item i is ordered, item j is also likely to be ordered in the same context. Type 2) DCs can be mined from historical data. Type 1) DCs may come from expert and engineering data, knowledge and simulations. As shown in Fig. 1, both types of demand change have to be fused into a stable new demand for all the items.

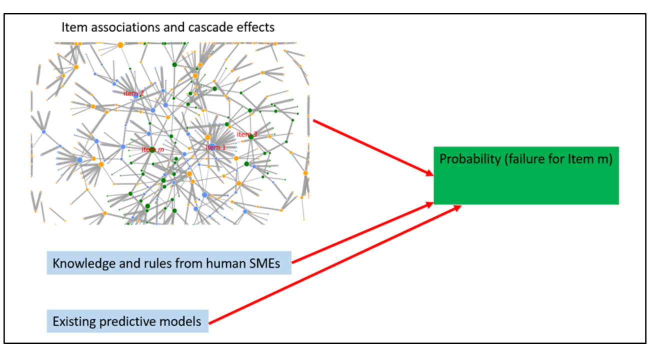

Fig. 1. LLA and probability of fail

In Eq. 4, $Assoc_{ij}$ measures how strong item i and j are demanded together. Probability and lift are the two measures defined in Eq. (1) and Eq. (3) in LLA to measure the strength of an association.

$$prob_{ij} = \frac{demand\ of\ item\ i, item\ j\ together}{demand\ of\ item\ j} \tag{1}$$

$$prob_i = \frac{demand\ of\ item\ i}{all\ item\ demands} \tag{2}$$

$$lift_{ij} = \frac{prob_{ij}}{prob_i} \tag{3}$$

$$DC_j^{(t+1)} = \sum_{i=1}^{N} Assoc_{ij} * DC_i^{(t)} \tag{4}$$

In this paper, LLA is used to compute the association network and graph, $prob_{ij}$, and $lift_{ij}$ from historical data. When a new operation condition occurs (Fig. 2 1)) that causes a demand change DC_i for item i. DC_i propagates through the association network to affect the whole system.

A. Game-theoretic Framework of LLA

It is interesting to note that there is a game-theoretic framework of LLA. If one item has a demand perturbation, the total system can be stabilized using a game-theoretic LLA as follows: Let the demand change vector for all item $\vec{x}$ for a new condition that starts with an initial perturbation of all the items in Eq. (5):

$$\vec{x}^0 = \begin{bmatrix} DC_1^0 \\ DC_2^0 \\ \vdots \\ DC_i^0 \\ DC_N^0 \end{bmatrix} \tag{5}$$

And

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \tag{6}$$

is the changed demand for the whole complex system. According to the game theory, a fixed point or a Nash equilibrium [12]

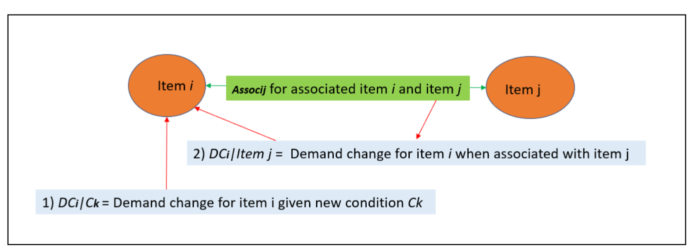

Fig. 2. Total demand change (TDC) caused by new conditions and associations

is achieved in Eq. (7):

$$\mathbf{A}\vec{x}^{*} = \vec{x}^{*}, \tag{7}$$

where $\mathbf{A} = \frac{1}{\lambda_{max}}\mathbf{Assoc}$ and $Assoc_{ij}$ is the $lift_{ij}$ matrix computed from LLA in Eq. (3). The stabilized demand $\vec{x}^{*}$ is the eigenvector of the maximum absolute eigenvalue of $\mathbf{A}$ [11], [13]. λ_{max} is the maximum eigenvalue of $Assoc$ and the maximum eigenvalue for $\mathbf{A}$ is 1. The demand change $\vec{x}^{0}$ propagates to the network of items. The final demand for all the items can be expressed as in Eq. (8):

$$\vec{x}^{fixed} = \sqrt{(DC_1^0)^2 + (DC_2^0)^2 + ... + (DC_N^0)^2} \begin{bmatrix} x_1^{*}2 \\ x_2^{*}2 \\ \vdots \\ x_i^{*}2 \\ x_N^{*}2 \end{bmatrix} \tag{8}$$

V. Use Case

A. Initial Data and Pre-processing

Currently in the USMC maintenance and supply system, an equipment or a part of an equipment fails, a service ticket is opened. A service ticket is then been taken care of in a long chain of actions such as to be replaced using spare parts, or to be repaired in various echelons, or to be requisitioned from the supply system. If an equipment or related parts can not be repaired, they are ordered or requisitioned. The service ticket is closed only when the equipment and parts are totally ready and back to use again. Fig. 3 shows an example of a service ticket. The time (usually days) between the open date and close date of the service ticket indicates how long does it take to make an equipment and associated parts ready to use again through the maintenance and supply chain. An USMC equipment is labeled using a table of authorized material control number (TAMCN). A TAMCN is associated with a set of parts specified using national stock number (RNSN).

To show the feasibility, we first fused data for a TAMCN from a wide range of databases. The days between opened and closed date for a service ticket is a derived attribute to measure the effectiveness of the USMC maintenance and supply chain. To look into how the days between opened and closed date correlate with other attributes, we aggregated the data to the service ticket number level with attributes related to 1) the maintenance data such as service number, service request type, defect code, operational status, echelon of maintenance, master priority code, count of job status dates, count of service cross-references, count of service parts, count of service activities, count of task numbers, and 2) the supply data such as count of RNSN, sum of part charge, count of document numbers, count of last parts update dates, count of requirement numbers, count of unit issue, count of item types, count of supply route locations, and 3) the equipment usage data such as owner unit address code, equipment operation time code, and meter reading. They are all potentially correlate with the days between opened and closed date. The sample data set contains 2065 service numbers/tickets and 599 (29%) of 2065 have the days between opened and closed date more than 65 days (65 days is the mean of the days between the open and close dates for the data set).

SR_NUMBER	TAMCN	SERVICE_REQUEST_TYPE	DEFECT_CODE	OPERATIONAL_STATUS	ECHELON	MASTER_PRIORITY_CODE	OWNER_UNIT	DATE_CLOSED	OPENED_DATE
11280122	E09497M	Maintenance - CM	PWRT.MDRV	Operational - Degraded	2	06 B-Urgent	M11700	06-SEP-19	10-MAR-14
11280122	E09497M	Maintenance - CM	PWRT.MDRV	Operational - Degraded	2	06 B-Urgent	M11700	06-SEP-19	10-MAR-14
11280122	E09497M	Maintenance - CM	PWRT.MDRV	Operational - Degraded	2	06 B-Urgent	M11700	06-SEP-19	10-MAR-14
11280122	E09497M	Maintenance - CM	PWRT.MDRV	Operational - Degraded	2	06 B-Urgent	M11700	06-SEP-19	10-MAR-14
11280122	E09497M	Maintenance - CM	PWRT.MDRV	Operational - Degraded	2	06 B-Urgent	M11700	06-SEP-19	10-MAR-14

Fig. 3. Initial maintenance data

B. Supervised Machine Learning and Predictive Modeling

In this section, we show how to apply supervised machine learning and predictive algorithms to predict, for a service ticket, the probability of the days between opened and closed date more than 65 days, based on other attributes. We use a data mining and machine learning open source tool Orange [5]. Orange consists of a wide range of data mining and machine learning algorithms including predictive models such as logistic regression, decision trees, naïve Bayes, random forest, and neural networks.

A typical procedure in predictive modeling or supervised machine learning is cross validation, which splits the data into k folds and uses k–1 folds for training and the remaining fold for testing ($k = 20$ in for our data). This procedure is repeated, so that each fold has been used for testing exactly once. The performance of the models is shown in Fig. 4. Among the various performance measures, an ROC curve is a graph showing the performance of a model at all prediction thresholds. This curve plots two parameters true positive rate (i.e., recall) as y-axis and false positive rate (x-axis). The area under ROC (AUC) provides an aggregate measure of performance across all possible prediction thresholds. A perfect predictive model's AUC is 1. The classification accuracy (CA) measures correctly predicted cases for all the classes if the predicted class has the highest predicted probability.

The precision is the fraction of the instances which are indeed positive out of those that are predicted positive. The recall is the fraction of the instances which are predicted positive out of those that are indeed positive. The F1 score is the harmonic mean of precision and recall. The Random forest has the best F1 score as shown in Fig. 4 for the data set compared to the other predictive models.

The confusion matrix gives the number/proportion of cases between the predicted and actual class as shown for the predictive decision trees in Fig. 5.

Furthermore, for supervised machine learning and predictive models, a lift curve shows the relation between the number of cases (i.e., service tickets) which are predicted positive (i.e., a service ticket closes more than 65 days) and those that are indeed positive and thus measures the performance of a chosen classifier against a random classifier (the straight line in Fig. 6. The graph is constructed with the cumulative number of cases in descending order of the predicted scores (i.e.,the predicted probability of a service ticket closes more than 65 days) on the x-axis and the cumulative number of true positives on the y-axis (recall). Lift curve allows to prioritize the cases based on the predicted scores. Fig. 6 shows lift curves for the predictive models (i.e., logistic regression, decision trees, naïve Bayes, random forest, and neural networks) tested for the data set. Logistic regression has the best lift curve compared to the other predictive models in Fig. 6.

While many supervised machine learning and predictive algorithms such as neural networks are criticized being black boxes and not explainable, some do provide better explainable AI (XAI [10]) features, for example, Fig. 7 shows a visualization of the predictive decision trees with rules specified in each leaf of the tree. An example rule reads "If $count\_requirement\_number > 0$ and $count\_last\_update\_date\_parts > 3$, then 78% (199 out of 255) of the service tickets close more than 65 days." The rule has a statistical significance $p < 0.0002$ as seen when comparing with the rate 29% (599 of 2065) in the population as a whole (random guess), the lift is 2.69.

Method	AUC	CA	F1	Precision	Recall
kNN	0.679	0.716	0.421	0.516	0.356
Tree	0.683	0.760	0.554	0.600	0.514
Random Forest	0.849	0.815	0.658	0.711	0.613
Neural Network	0.848	0.802	0.634	0.683	0.593
Naive Bayes	0.828	0.713	0.618	0.503	0.800
Logistic Regression	0.859	0.822	0.651	0.756	0.573

Fig. 4. Orange prediction accuracy for predictive algorithms using cross-validation $k = 20$

Actual \ Predicted	days_bt_opened_closed_less_65	days_bt_opened_closed_more_65	Σ
days_bt_opened_closed_less_65	81.2 %	40.0 %	1466
days_bt_opened_closed_more_65	18.8 %	60.0 %	599
Σ	1552	513	2065

Fig. 5. Orange confusion matrix for the decision trees predictive algorithm

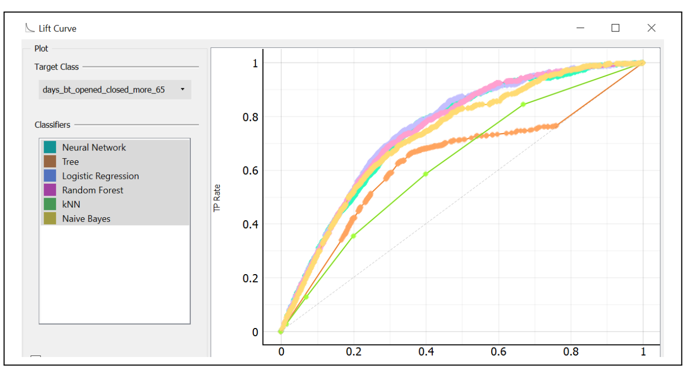

Fig. 6. Orange lift curve shows

C. data-mining Association Patterns and Networks Using LLA

Next, we applied LLA to data mine associations among the part failures within the same contexts, i.e., the same open dates. As shown in Fig. 10, we first used LLA to group all the parts (labeled by the RNSN numbers and NOMENCLATURE) based on the service ticket number. We also attached information that if a ticket closed more than 65 days and if there is an indication of spare parts used for the ticket. Each row represents all parts needed within the same ticket. There are total 1240 tickets in the data set with the RNSN numbers and NOMENCLATURE. In other words, 60% (1240/2065) of service tickets required requisitions of parts.

We applied LLA to compute pair-wise parts associations, i.e., the parts that are needed/ordered/demanded in the same service ticket. Associated spare parts might be stockpiled in the same manner should one fail suddenly in a new and disrupted condition.

As shown in Fig. 8, LLA generated so-called themes – Each theme is a cluster of parts that appeared together in more than one service ticket. The association strength is measured by the

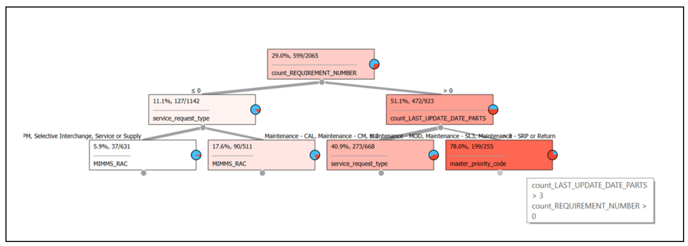

Fig. 7. Decision tree rules visualization shows how an example of explainable AI

ID	Theme Keywords
12(P)	4730015331439_coupling_half_quick 4940003335541_gun_air_blow,4730015331439_coupling_half_quick 4930002881511_adapter_grease_gun,4930002881511_adapter_grease_gun 4730015331439_coupling_half_quick,4940003335541_gun_air_blow 4930002881511_adapter_grease_gun,4930002881511_adapter_grease_gun 4940003335541_gun_air_blow,4940003335541_gun_air_blow 4730015331439_coupling_half_quick
21(A)	2910010228183_filter_element_flui 2940013141345_filter_element_flui,2910013538540_filter_element_flui 2940013141345_filter_element_flui,2940013141345_filter_element_flui 2910013538540_filter_element_flui,2910010228183_filter_element_flui 2910013538540_filter_element_flui,2940013141345_filter_element_flui 2910010228183_filter_element_flui,2910013538540_filter_element_flui 2910010228183_filter_element_flui
6(A)	5305011580835_screw_cap_hexagon_h 5360011791219_spring_helical_comp,5305011580835_screw_cap_hexagon_h 2540218789053_mirror_rearview,2540218789053_mirror_rearview 5360011791219_spring_helical_comp,5360011791219_spring_helical_comp 5305011580835_screw_cap_hexagon_h,5360011791219_spring_helical_comp 2540218789053_mirror_rearview
13(A)	5120010798024_wrench_box_and_open 5120013491438_wrench_box_and_open,5120013491438_wrench_box_and_open 5120010464951_wrench_box_and_open,5120013491438_wrench_box_and_open 5120010798024_wrench_box_and_open,5120010464951_wrench_box_and_open 5120010798024_wrench_box_and_open

Fig. 8. Themes discovered by LLA

"lift" measure (similar to the meaning of the lift curve in the predictive models in Orange) described in Equation (3).

Fig. 9 shows associated parts in Group 12. Part $5120002370977\_socket\_socket\_wrenc$ and $8415010920039\_mitten\_heat\_protect$ have a link strength $lift = 82.7$ because they were serviced 4 times together while $8415010920039\_mitten\_heat\_protect$ was serviced 5 in total and $5120002370977\_socket\_socket\_wrenc$ was serviced 12 times in total among a total service number of 1240, therefore, the lift is 82.7=(4/5)/(12/1240). In other words, when $8415010920039\_mitten\_heat\_protect$ was serviced (i.e., ordered or requisitioned), 80% of the time $5120002370977\_socket\_socket\_wrenc$ was also serviced, which is 82.7 times of $5120002370977\_socket\_socket\_wrenc$'s service in the whole data set (12 of 1240). Total service demand of $5120002370977\_socket\_socket\_wrenc$ may be caused by services of all other associated items as shown. Should an item's demand changes in a new operation condition, associated items may demand even more in the new condition.

Fig. 11 shows a ground truth POD computed using the number of service tickets demanded for a part (e.g., 12 for $5120002370977\_socket\_socket\_wrenc$) divided by the total service tickets (i.e.,1240). We computed Eq. (7) and Eq. (8), and the correlation to POD is plotted in Fig. 11. The Pearson correlation is 0.13 (p=0.025). The centrality measure of total degree [6], [11] computed from LLA has the correlation 0.32 ($p < 0.0001$) with POD as plotted in Fig. 12. Both correlations are statistically significant, while the total degree measure has a higher correlation. Fig. 13 shows an example of predicted POD sorted based on either of the two scores. In a summary,

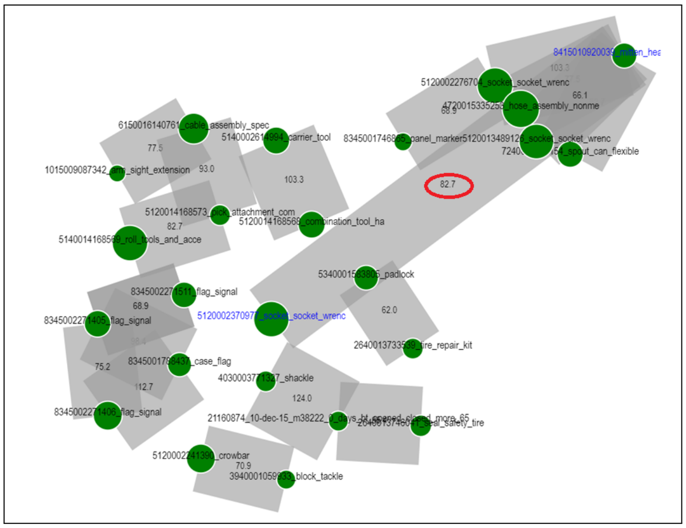

Fig. 9. LLA Group 12 lift: Associated parts in Group 12. Part $5120002370977\_socket\_socket\_wrenc$ and $8415010920039\_mitten\_heat\_protect$ have a link strength $lift = 82.7$ because they were serviced 4 times together while $8415010920039\_mitten\_heat\_protect$ was serviced 5 in total and $5120002370977\_socket\_socket\_wrenc$ was serviced 12 times in total among a total service number of 1240, therefore, the lift is 82.7=(4/5)/(12/1240). In other words, when $8415010920039\_mitten\_heat\_protect$ was serviced (i.e., ordered or requisitioned), 80% of the time $5120002370977\_socket\_socket\_wrenc$ was also serviced, which is 82.7 times of $5120002370977\_socket\_socket\_wrenc$'s service in the whole data set (12 of 1240).

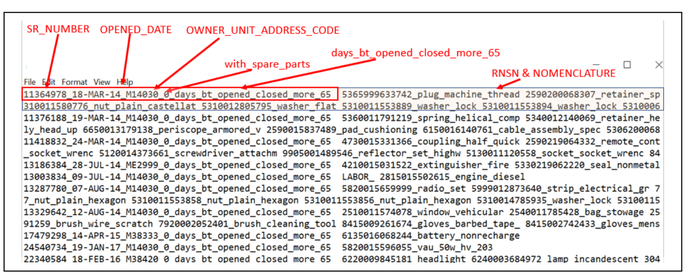

Fig. 10. LLA association input data

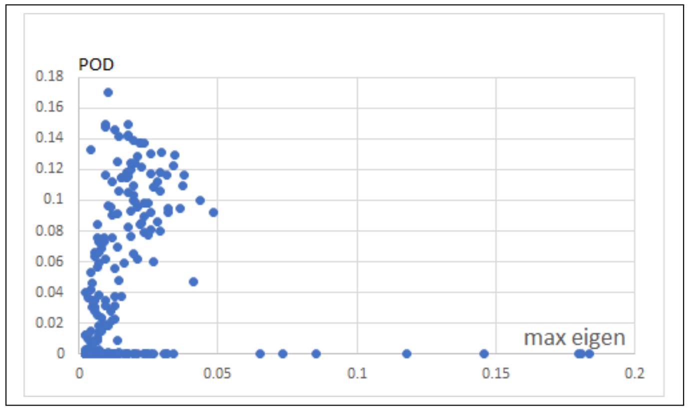

Fig. 11. POD sorted using max eigenvector scores

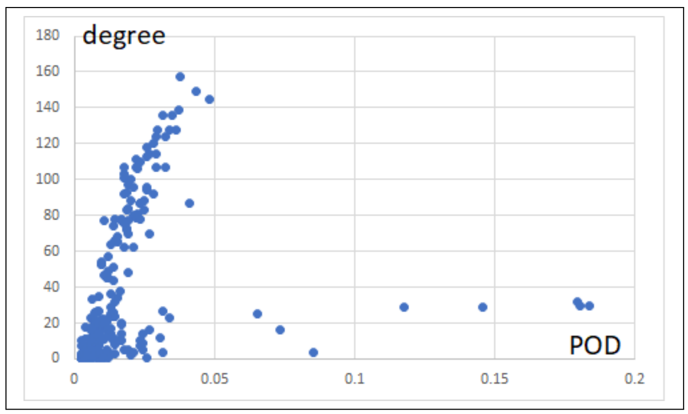

Fig. 12. Correlation between POD and total degree scores from the LLA parts association network

RNSN	# services	total services	POD	LLA group	degree	max eigen
4940003335541_gun_air_blow	47	1240	0.037903	12	157	0.116376044
4930002881511_adapter_grease_gun	54	1240	0.043548	12	149	0.099785756
4730015331439_coupling_half_quick	60	1240	0.048387	12	145	0.092177193
5120002405364_handle_socket_wrenc	46	1240	0.037097	12	139	0.10915423
5120010458565_socket_socket_wrenc	43	1240	0.034677	12	136	0.129753311
5120014373661_screwdriver_attachm	39	1240	0.031452	12	136	0.116809139
5120010246144_socket_socket_wrenc	37	1240	0.029839	12	128	0.130939631
5120010246146_socket_socket_wrenc	42	1240	0.033871	12	128	0.122546367
4210015194385_extinguisher_fire	45	1240	0.03629	12	128	0.094424531
5120010798024_wrench_box_and_open	36	1240	0.029032	12	124	0.118521432
5995011608030_cable_assembly_spec	40	1240	0.032258	12	124	0.092004679

Fig. 13. Eigenvalue-sorted items

the association network and derived attributes such as total degree centrality can be used to predict POD, POF, infer needed spare parts, and reduce maintenance and supply delays together with other attributes and databases.

VI. Conclusion

We showed a use case for a generic framework of LAILOW in detail in the context of the USMC maintenance and supply big data and complex system. We used LLA to compute stable and sensitive components of a complex system with respective to a perturbation. LLA allows discover and search for item associations that can be used to improve predicting the demand, prepare spare parts, and improve operational availability and readiness. The future work is to test the LAILOW framework for more database integration (e.g., equipment movement and usage, manpower, engineering, and transportation databases), more core equipment of the USMC, and perform simulations/games of new operation conditions and cascade effects.

VII. ACKNOWLEDGMENTS

Authors would like to thank the Office of Naval Research (ONR)'s Naval Enterprise Partnership Teaming with Universities for National Excellence (NEPTUNE 2.0) program and the Naval Postgraduate School (NPS)'s Naval Research Program (NRP) for supporting the research. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied of the U.S. Government.

References

[1] Major Mata, Gabe (2019). Machine learning problem statement sustaining the force. USMC Report.

[2] Major Mata, Gabe (2019). Optimize logistics performance through ML. USMC Report.

[3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B.; Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In Advances in Neural Information Processing Systems, 2672–2680.

[4] Arora, S.,Ge, R.,Liang, Y.,Ma, T., and Zhang, Y. (2017). Generalization and equilibrium in generative adversarial nets (GANs). arXiv preprint arXiv:1703.00573.

[5] Orange, (2020). Retrieved from https://orange.biolab.si/

[6] Girvan, M., and Newman, M. E. J. (2002). Community structure in social and biological networks. Proceedings of the National Academy of Sciences, USA, 99(12), 7821–7826.

[7] Freeman, L.C. (1979). Centrality in social networks I: conceptual clarification. Social Networks, 1: 215-239.

[8] Zhao, Y., Mackinnon, D. J., Gallup, S. P. (2015). Big data and deep learning for understanding DoD data. Journal of Defense Software Engineering, Special Issue: Data Mining and Metrics.

[9] US patent 8,903,756, (2014). System and method for knowledge pattern search from networked agents. Retrieved from https://www.google.com/patents/US8903756

[10] DARPA XAI 2018. https://www.darpa.mil/program/explainable-artificial-intelligence. DARPA.

[11] Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E, vol. 74, no. 3.

[12] Nash, J. (1950). Equilibrium points in n-person games. Proceedings of the National Academy of Sciences 36(1):48-49.

[13] Jackson, M. and Zenou Y. (2014). Games on networks. Retrieved from https://web.stanford.edu/ jacksonm/GamesNetworks.pdf.