Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

THE ACTIVITY-ENTITY-IMPACT METHOD: UNDERSTANDING BOTTLENECK BEHAVIOR OF SIMULATION MODELS DEMONSTRATED BY AN EMERGENCY DEPARTMENT MODEL

Nikolaus Furian
Clemens Gutschi
Dietmar Neubacher

Institute of Engineering and Business Informatics
TU Graz
Kopernikusgasse 24
Graz, 8010, AUSTRIA

Cameron Walker
Michael O'Sullivan

Department of Engineering Science
University of Auckland
70 Symonds Street
Auckland, 1010, NEW ZEALAND

ABSTRACT

Simulation models are often used to gain a better understanding of a system's sensitivity to changes in the input parameters. Data gathered during simulation runs is aggregated to Key Performance Indicators (KPIs) that allow one to assess a model's or system's performance. KPIs do not provide a deeper understanding of the causes of the observed output because this is not their primary objective. By contrast, dynamic bottleneck methods both identify elements that yield the largest gain in productivity with increased availability and also visualize these elements over time to enable bottlenecks to be better understood. In this paper we discuss whether dynamic bottleneck detection methods can be utilized to identify, measure, and visualize causes of observed behavior in complex models. We extend standard bottleneck detection methods, and introduce the Activity-Entity-Impact-Method. The practicality of the method is demonstrated by an example model of a typical Emergency Department setting.

1 INTRODUCTION

A Bottleneck (BN) of a system is defined as the element for which an increase in availability or productivity yields the largest gain in terms of system output for a given period of time (Chang et al. 2007). Dynamic BN detection methods use mathematical models or algorithms to determine such BNs over time. Most BNs are identified by using data gathered from the real system and then feeding it into the detection methods (i.e., mathematical methods or algorithms). On the other hand, simulation models are also frequently used to investigate the behavior of real systems, and their sensitivity to parameter changes. Hence, they are also used to detect BN elements and assess the impact of increasing or reducing the availability and productivity of such elements. Further, it is often implied that simulation models are useful tools to understand the dynamics of systems. However, understanding system dynamics and identifying dynamic BNs requires understanding and analysis of the behavior of the simulation model itself. The definition of KPIs plays an important role in gaining a better understanding of simulation models and/or systems. KPIs allow the user to assess the impact of changes to the input parameters on the performance of the system or individual elements of it. However, recent literature shows that commonly used KPIs do not capture the dynamic relations of entities within and impact on models/systems (Barrera-Diaz et al. 2018; Furian et al. 2018). In particular, KPIs are often computed by rigorously aggregating measures over time, e.g., the *average resource utilization* or *throughput*, but do not account for interactions between elements in the system. This may lead to the practice where simulation models are used as black boxes to assess the sensitivity of a system, but little support is provided by standard KPIs to understand causes of certain behavior. It has to be noted that this is not the main purpose of KPIs, but instead tools to gain a better understanding of a model's dynamic behavior may be of interest.

In this paper we investigate whether dynamic BN detection methods are suitable to gain a deeper understanding of not only a real system's, but also a model's behavior. We assess if it is possible to detect possible causes for, and gain knowledge by evaluating, the shift of a critical elements over time. First, we apply a well-established dynamic BN method, the *Active Period Method (APM)* on an example simulation model and discuss results. Based on this discussion, we introduce a new simple BN detection and impact method, called the *Activity-Entity-Impact-Method (AEIM)*, that is based on the main ideas of APM. Similar to APM, the proposed approach assesses elements by distinguishing between active and inactive periods, where inactivity is caused by blocking or starvation. However, in addition to evaluating the length of blocked periods, AEIM attempts to identify causes for such blocking states. As APM originates from the manufacturing industry, where often production sequences are rather rigid and causes of blocking are obvious, its main focus is not the identification of these causes. For more complex systems, where resources take part in a variety of activities, possibly in different combinations, the cause of inactive periods may be less apparent. Using the example of an Emergency Department (ED) model, we demonstrate the usefulness of AEIM to not only detect such causes, but also to measure them by various additional KPIs. However, like APM, the method itself is domain independent and applicable to other fields such as modular production, logistics, construction, project management, etc.

The paper is structured as follows: In the following section we provide a brief review of relevant work on BN detection methods and simulation KPIs. Section 3 introduces the example ED model used throughout this paper. In section 4 we apply the APM to the ED model and discuss results and insights gained. In section 5 the AEIM is presented and its application to the ED model is outlined in section 6. The paper concludes with final remarks and suggestions for further research.

2 BACKGROUND

In this section we give a brief overview of theories and methods that are relevant for this paper. First, we revisit established dynamic BN detection methods. Second, common KPIs used to analyze models in the fields of manufacturing and ED simulation are introduced. For general information on statistical analysis for simulation modeling the reader is referred to Law (2015) and Currie and Cheng (2016).

2.1 Dynamic Bottleneck Detection Methods

Bottleneck detection methods have been well studied in the production industry. In addition to, standard measures, such as queue length and resource utilization, dynamic methods have gained significant attention. These include, among others, the *Active Period Method* (Roser et al. 2001), the *Average Active Period Method* (Roser et al. 2003), *Blocking and Starving Probabilities* (Kuo et al. 1996), and *Shifting Bottleneck Detection* (Roser et al. 2003). For comparisons and reviews of these methods the reader is referred to Roser and Nakano (2015) and Wang et al. (2005). As the *Active Period Method* and *Shifting Bottleneck Detection* form the basis of this paper, they are outlined in the next section and a discussion of their applicability for more complex-systems follows in section 2.1.2.

2.1.1 The Active Period Method and Shifting Bottleneck Detection

The main idea underlying these methods is that the state of entities, e.g., machines or other resources, can be classified into active and inactive periods. For production machines active states include, for example, production, repair, and tool change, whereas inactive periods include blocking (e.g., a full downstream buffer) and starving (e.g., an empty upstream buffer) states. For a given point in time the machine with the longest current active period is denoted as the dynamic BN. However, it is assigned the BN status not only for this particular point in time, but also for the entire active period that it is currently in. Hence, if two active periods overlap, without one being covered entirely by the other, both machines are denoted as BNs (as both will be the longest active elements at some point). The *Shifting Bottleneck Detection Method (SBDM)* then further classifies each machine by whether they are the sole BN or part of a shifting BN (one of multiple machines with overlapping active periods) and reports the proportion of each state over time.

2.1.2 Discussion of Bottleneck Methods

The APM and SBDM have been shown to detect dynamic BNs correctly for production systems. Although, Roser et al. (2003) state that they can be applied to any system that allows the classification of active and inactive states, the application to more complex-systems, e.g., flexible and modular production systems or EDs, raises some questions.

In (standard) production lines or systems, the coupling between tasks and resources is usually more rigid than in other systems. In other words, a single resource is often only responsible for one task, e.g., milling. In more complex-systems, e.g., an ED, resources perform multiple tasks, often in combination with other resources. While this may be easily addressed by the APM, as long as active and inactive states can be distinct, it can also be observed that blocking and starving of resources may occur for various reasons. In production lines, blocking and starving occurs with respect to full or empty buffers. Blocking in a complex system may additionally occur due to other resources being active or blocked. While this may not be relevant for identifying which resource is the bottleneck in a specific situation, measuring active periods only provides limited information on why this resource is crucial. Further, considering the example of an ED, there may be multiple instances of a resource type present, e.g., more than one doctor. Therefore, it seems more important to determine which type of resource is currently the BN, and not which instance of it. Some of these issues will be further discussed in section 4.

2.2 Simulation Key Performance Indicators

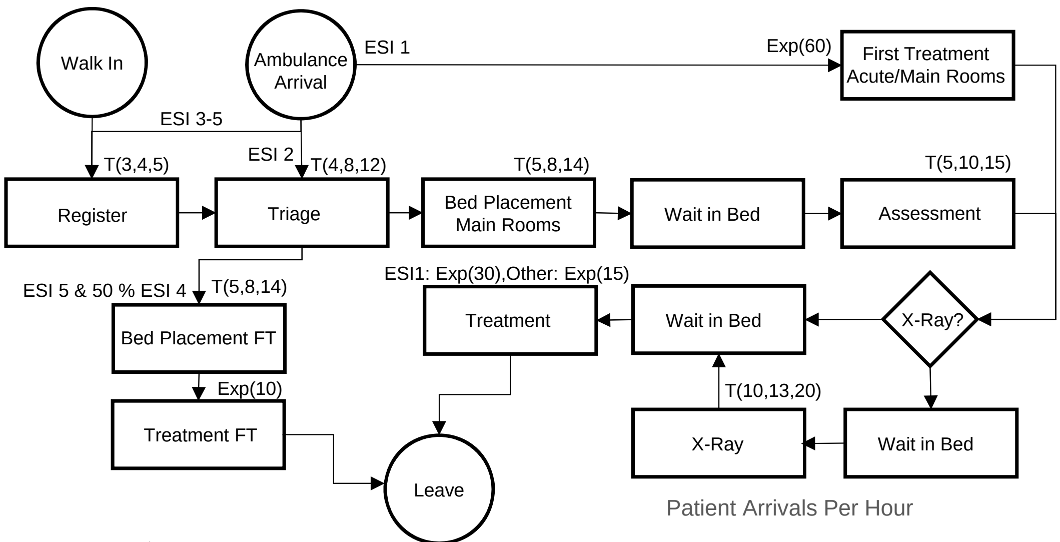
KPIs are measures that reflect the performance of a system over a given period of time; see for example Parmenter (2007). In the context of simulation models they are often used as tools to assess and investigate the behavior of the model. In a previous study, a wide range of ED models were analyzed with regard to a variety of modeling and simulation aspects, including KPIs used. Thereby, it became apparent that most commonly used measures aggregate data gathered from the models into a single quantity. Most prevailing examples include the *Length of Stay (LOS)*, *Time to Doctor (TTD)*, *Resource Utilization (RU)*, *average waiting times*, or *average queue lengths*. For an exhaustive list of KPIs used, the reader is referred to Furian et al. (2018). KPIs used in other fields follow similar structures. In the manufacturing industry they commonly include *work in progress*, *throughput*, *resource utilization*, *lead times*, or *equipment efficiency*; see for example Barrera-Diaz et al. (2018).

3 THE EMERGENCY DEPARTMENT MODEL

In this section we introduce the model that serves as a basis for demonstrating the principles and use of the AEIM method. The model is based on previous work of the authors, where they analyzed and investigated a large number of ED models (Furian et al. 2018).

3.1 Patient Categories, Arrival Times, and Pathways

Analogue to many published models, patient categories are mainly defined over triage grades by the ESI system: ESI1 denotes patients that require immediate life-saving interventions; ESI2 represents high risk patients that should be seen immediately; ESI3 includes patients that require many resources or show potentially threatening vital signs; ESI4 patients require one and ESI5 no resource. Further, patients are classified by the mode of arrival (ambulance or walk-in). However, in accordance with, for example, Day et al. (2013) here we assume that ESI5 and 50% of ESI4 patients are eligible for a Fast Track (FT) process through the ED. The distributions of patients over ESI grades (based on average values of models in Furian et al. (2018)), mode of arrival (based on Abo-Hamad and Arisha (2013)) and share of patients that require X-Ray services (based on Gunal and Pidd (2006)) are summarized by Figure 1. Patient arrival rates (per hour of the day) are taken from Lim et al. (2013) and also reported by Figure 1.



ESI Grade	1	2	3	4	5
Share	2.7%	18.0%	47.3%	25.8%	6.1%
By Ambulance	96.0%	60.0%	39.0%	19.0%	28.0%
Need X-Ray	74.0%	62.0%	53.0%	37.0%	31.0%

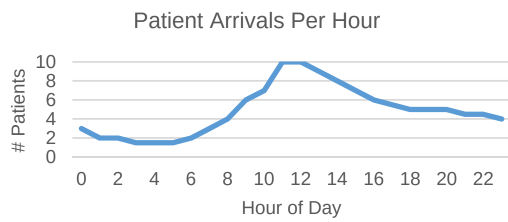


Figure 1: Patient processes and data. Note that the duration of activities are either triangular (T(min, mode, max)), or exponentially (Exp(mean)) distributed. Further, ESI denotes triage grades and FT represents the Fast Track process.

The main patient path through the model follows common structures observed by Furian et al. (2018). Upon arrival, patients register, get triaged, and are then placed in a bed by a nurse. This *Bed Placement* activity may include preparation tasks, such as taking a blood sample. This is followed by an *Assessment* activity performed by a doctor, possibly an X-Ray examination (other diagnostic tests are omitted for simplicity reasons) and the final *Treatment* activity where both a doctor and a nurse are present. ESI1 and ESI2 patients that are arriving via ambulance skip *Register* and *Triage* and further, ESI1 patients undergo a *First Treatment* with both a nurse and a doctor present instead of the *Assessment* (in order to stabilize the patient). For the different processes, durations, and other patient data see Figure 1.

3.2 Staffing and Physical Resources

The staff types within the model include: doctors, general nurses, triage nurses, register nurses (or clerks), diagnostic personnel and specially trained nurses for FT procedures (equivalent to mid level providers, see Furian et al. (2018)). We assume that three doctors, two general nurses, and one of each of the other staff type are present during relevant hours of the day. Note that for simplicity reasons and since we analyze the behavior model only during day hours, we did not include more realistic shift patterns that account for the varying arrival rates of patients during the day. Physical resources include a registration desk, a triage room, main ED rooms, acute rooms, a FT room and a X-Ray facility. Besides five main ED rooms, each resource is represented once in the model, e.g. there is one acute room.

3.3 Control Policies

The policies to control the logic also follow common principles for ED models. Dispatching of activities is performed in the following order: *First Treatment ESI1*, *Bed Placement*, *Assessment* and *Treatment*. This prioritizes new patients over old patients as witnessed regularly among ED models, see for example Ghanes et al. (2014). *FT*, *Triage*, *X-Ray* and *Register* activities are dispatched separately. Patients are selected by triage grades and FIFO for breaking ties. Patients see the same doctor during their entire stay and occupy their room for the entire stay (even while being at X-Ray). Resources are selected with respect to idle times to ensure an evenly distributed workload. Whenever a FT nurse is available and no FT patients are present, he/she assists in the main ED. In principle, ESI1 patients are assigned the acute room, however, if it is occupied they are assigned to a free main room or whichever becomes free first.

4 DYNAMIC BOTTLENECK DETECTION TO ANALYZE SIMULATION MODELS

One of the main goals of this paper is to investigate whether BN detection methods can be used to analyze simulation models. However, raising that question, one has to be aware that simulation models themselves have been used numerous times to identify BNs via evaluating *what-if* scenarios on the availability and productivity of entities. Hence, the sole identification of BNs may seem redundant at first glance. However, as the APM and SBDM enable the analysis of behavior over time, they may provide insight regarding the dynamics that cause entities to become BNs, especially because simulation models are often used in a black box fashion when evaluating *what-if* scenarios.

As mentioned in section 2.1.2, some difficulties arise when applying the APM to models like the ED model stated in the previous section. First, the APM does not define how to aggregate multiple instances of the same resource type, e.g. doctors, to a single resource type that may act as a BN. Experiments with the model have shown that counting discretized time intervals (e.g. per minute) where at least one resource of a type acts as a BN (shifting or sole) is one possible approach to overcome this issue (given that the workload of entities of the same type is somehow balanced). Second, while the definition of active and inactive states is straight forward for human resources in the model, assigning those states to rooms raises some questions. It is apparent that rooms are active during patient related activities. However, during *Wait In Bed* or when patients are currently at diagnostics those resources are occupied but not actively used. Experiments have shown that classifying those states as inactive leads to correct identification of BNs. This makes also sense from a theoretical point of view as, during those times, rooms would be able to be active but are not used due to other entities not being available.

Applying APM to the ED model of the previous section yields the following BN and resource utilization results illustrated in Table 1. Results include simulation runs of the base scenario (D3-N3-M5 = 3 doctors, 3 nurses, 5 main rooms) and scenarios with an additional resource of each type.

Note that for this evaluation we aggregated general nurses and FT nurses into one resource type. Results show that APM with outlined adaptions identifies BNs correctly for this model. However, this information can also be gained by experimenting with the simulation model itself. In order to investigate dynamic BNs, Figure 2 illustrates the BN behavior of the main entities in the model and the arrival of patients over time.

One may observe that BNs are shifting in the model. However, the most dominant behavior is that more than one resource type is acting as a BN. Moreover, no apparent correlation in the pattern of patient arrivals and BN behavior can be observed. It is not apparent from this analysis what behavior in the model is the major cause of blocking and/or starving. As APM focuses on entities/resources and their states it does not provide a deeper understanding of the dynamics of models and/or systems. An appropriate method should put more emphasis on the relationship between entities, activities and queues. Hence, in the following section we extend the APM method to allow a more detailed investigation of complex models and/or systems.

Table 1: Average BN results using APM over 200 simulation runs over two days, where the first day is used as a warm up-phase. Scenarios Dx-Ny-Mz are defined over available resources (i.e. x-doctors, y-nurses, z-main rooms). BN denotes the number of minutes at least one resource of the specific type was BN, RU the resource utilization of the resource between 7am and 10pm of the second day.

	Doctors		Nurses		Main Rooms		KPIs	
Scenarios	BN	RU	BN	RU	BN	RU	LOS	TTD
D3-N3-M5	234.51	0.61	520.52	0.71	450.62	0.47	162.05	116.10
D4-N3-M5	139.89	0.48	550.22	0.73	433.56	0.48	142.36	98.95
D3-N4-M5	347.56	0.65	341.62	0.56	488.24	0.50	130.00	88.44
D3-N3-M6	251.43	0.63	536.26	0.72	385.42	0.41	151.50	99.48

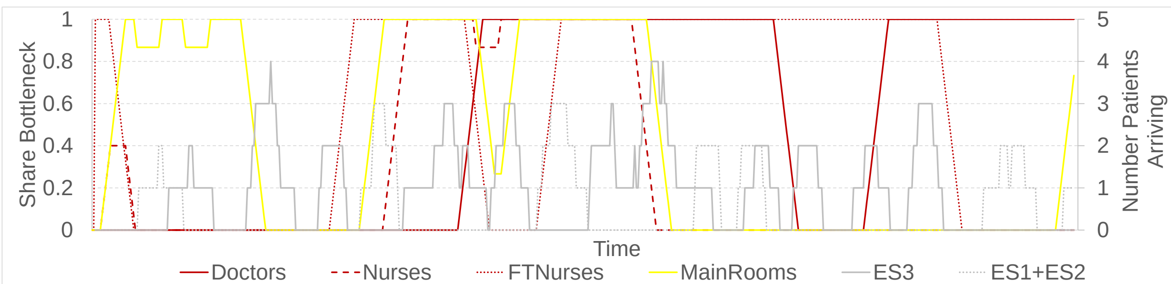


Figure 2: The dynamic BN behavior between 7am and 5pm for most relevant entities and patient arrivals of most relevant categories for a single run of the base scenario.

5 THE ACTIVITY-ENTITY-IMPACT METHOD

To put activities as the main focus of simulation model is not new, and has been published previously, in different forms, by (among others) Heath et al. (2011) and the authors of this paper. The Hierarchical Control Conceptual Modeling (HCCM) framework (Furian et al. 2015) uses a slightly different terminology and perspective to describe elements of a model or system. In particular, it extends standard queues by instead using requests and corresponding lists that represent entities waiting for specific requests (for an activity to take place) to be granted. Thereby, the rigid coupling of queues and resources is relaxed. While this paper is not based on the HCCM framework, we use its terminology and refer to queues, or waiting entities, as requests.

The main idea of the AEIM is to design KPIs that capture the dynamics resulting from inter-dependencies of entities, activities and requests, and make them visible in a simple and intuitive way.

5.1 The AEIM Matrix

For defining the components of the AEIM matrix, let in general $\{a_1,\dots,a_n\}$ be the the set of activity types and $\{e_1,\dots,e_m\}$ be the set of resource types. Further, we denote by $arf_{i,j}$ $\forall i \in [1,\dots,n], j \in [1,\dots,m]$ the *activity resource factor*, which is 1 if resource type e_j is required for activity a_i (else 0), and by $aif_{i,j}$ $\forall i, j \in [1,\dots,n]$ the *activity impact factor*, which is 1 if a_i and a_j share at least one resource type (else 0). Hence, $aif_{i,j}$ can be computed as $aif_{i,j} = \min\left(1, \sum_{k=1}^{m} arf_{i,k} \cdot arf_{j,k}\right)$. For a specific point in time, let $na(a_i)$ be the number of activities of type a_i currently performed, $nr(a_i)$ be the number of requests for activity of type a_i currently present, and $nae(e_j)$ be the number of active (non-idle) resources of type e_j.

We make some remarks on the definitions of $aif_{i,j}$ and $arf_{i,j}$. In some models a resource may be attached to an entity for a period of time that stretches over multiple activities. For example a bed may be assigned to a patient at the beginning of the stay in the ED and remain occupied regardless of whether the patient is in the bed during certain activities (e.g. diagnostics), or, in manufacturing, a workpiece carrier that is assigned to a part for multiple operations. In such cases $arf_{i,j}$ is set only to 1 if a_i is the activity that the resource is assigned to. Considering the example used in this paper, patients are assigned to rooms for their entire stay (except FT patients). Hence, the *Wait In Bed* activity shares the resource room with the activity *Bed Placement*, but not with any other activity. For activities that can use multiple resource types as one resource (e.g. the *Bed Placement* of non-FT patients can be performed by nurses as well as FT nurses) both activity resource factors are set to 1.

The AEIM matrix is composed of several elements. First, the activity impact sub-matrix, A^{aai} that reflects how activities impact each other in terms of execution and requests. In particular, if activities a_i and a_j compete for at least one resource of the same type, A^{aai} measures how the execution of a_i impacts the *waiting list* for activity a_j by multiplying associated parameters, see Table 2. The resource activity sub-matrix, A^{eai}, measures the impact of resource availability with respect to currently present requests (or in other words queue lengths). Therefore, a resource type is considered blocking if none of its instances are idle but there are requests filed that would require the participation of this resource type.

Vectors $\hat{a}$ and $\tilde{a}$ denote the total influence an activity is imposing on all other activities, and the total influence imposed on an activity by all others respectively. The vector *na* reports the number of activities per type currently in progress. The vector *nbe* denotes the number of resources of a specific type that are currently a sole or shifting bottle neck with respect to the APM method. Table 2 illustrates how the AEIM matrix is composed of the previously outlined elements and how individual elements are computed.

Table 2: The AEIM matrix and its components for a specific point in time.

	a_1 $\dots$ a_n	$\tilde{a}$	na	e_1 $\dots$ e_m
a_1 $\vdots$ a_n	$a^{aai}_{i,j} = nr(a_i)\cdot na(a_j)\cdot aif_{i,j}$	$\tilde{a}_j = \sum_j a^{aai}_{i,j}$	$na(a_i)$	$a^{eai}_{i,j} = nr(a_i)\cdot \lfloor nae(e_j)/ne(e_j)\rfloor \cdot arf_{i,j}$
$\hat{a}$	$\hat{a}_j = \sum_i a^{aai}_{i,j}$		nbe	$nbe(e_j)$

5.2 Aggregation over Time

The definition of the AEIM matrix enables the analysis of the behavior of a model over time or at a given point in time. However, it might be of interest to investigate the aggregate or average behavior for a chosen time span. Therefore, it is necessary to aggregate individual AEIM matrices. For the A^{aai} sub-matrix, vector $\hat{a}$ and vector $\tilde{a}$ aggregation over time can be derived by computing average values of individual elements. In particular, entries of the aggregated sub-matrix $\overline{a}^{aai}_{i,j}$ over time steps $t \in \{1,\dots,T\}$ are given by $1/T\sum_{t=1}^{T} a^{t,aai}_{i,j}$ (analogue for $\hat{a}$ and $\tilde{a}$). For *na* and *nbe* we count the time steps where at least one activity of a given type is active, or respectively at least one resource of a given type is BN. Hence, elements of the aggregated vector $\overline{na}^t$ for time stamps $t \in \{1,\dots,T\}$ are computed by $\overline{na}(a_i) = \sum_{t=1}^{T} \min(1, na^t(a_i))$ (analogue for *nbe*). In the same fashion, for the sub matrix A^{eai} we count the number of entries that are larger or equal to 1. This is motivated by the observation that the number of unfulfilled requests is already represented in the A^{aai} matrix and from an additional point of view the frequency of an resource showing a blocking behavior seems more informative than its magnitude.

6 APPLICATION AND RESULTS

In this section we demonstrate the use of the AEIM method by applying it to the model outlined in section 3. The model was implemented using HCDESLib, an open source discrete event simulation library in C# (Furian et al. 2016). Results are based on the average of 200 simulation runs, each simulating two days, using the first day as a warm-up period. Further, KPIs are computed on the results during 7am and 10pm of the second day, as those are the busy hours during daytime of the ED. The analysis of the model is performed in two stages. In section 6.1, activities of specific types, i.e. treatment, triage, etc., are investigated in a summarized way over all patient categories (in the model they are still differentiated). The impact of different triage grades is investigated in section 6.2.

6.1 Analyzing the core structures of the ED model

In order to apply the AEIM method, parameters $aif_{i,j}$ and $arf_{i,j}$ have to be defined for all pairs of activities and pairs of activities and entities. Note that the assistance of FT nurses in *Bed Placement* or *Treatment* activities for non FT patients may be treated different to the ones performed by standard nurses, as their impact on the system is different. To address this behavior and limit the number of activities we summarized those by an *Assist FT* activity to measure the impact of FT assistance. In particular, *Assist FT* has an impact on *Bed Placement FT* and *Treatment FT* but not vice-versa and not on other main ED activities.

First, looking at standard KPIs, see Table 3, we observe typical values and behavior of LOS and TTD. Measures are increasing for lower priority patients apart from FT patients and ESI1 patients (as those result in significant longer treatment times). Resource utilization of main entities show that doctors are utilized approximately 61%, nurses 68%, the FT nurse 76%, triage nurse 66%, register nurse 33%, x-ray room 55%, main rooms 47% (not counting waiting in bed times) and the acute room 23%. From these statistics it can be concluded that patients wait for a relatively large share of their stay to be placed in a room. However, possible reasons causing long waits for *Bed Placement* other than the workload of resources are not visible from standard KPIs. Hence, we investigate the AEIM matrix of a single simulation run over time. As the visualization of a matrix over time is not intuitive we report relevant entries (combinations of activities and activities and entities) over time using a resolution of one minute.

Table 3: Standard KPIs in minutes.

Category	Overall	ESI1	ESI2	ESI3	ESI4	FT
LOS (TTD)	162.1 (116.1)	76.5 (3.6)	67.4 (19.9)	157.9 (105.0)	421.0 (371.3)	89.5 (70.7)

Figure 3 shows some interesting behavior of the models. First, it seems that *Wait in Bed* has a significant influence on *Bed Placement* (significant large values in the A^{aai} matrix, leftmost red box). This means that there are often patients waiting for a room to be placed in while other patients are blocking rooms at the same time without receiving care. Accordingly, it can be observed that nurse (both standard and FT) availabilities are often not the cause for delay of *Bed Placement* (values in the A^{eai} matrix (middle red box) indicate that there are idle nurses while patients wait for bed placement), but rather blocked rooms by waiting patients ($na(a_i)$ values in the rightmost red box indicate that patients are waiting in their rooms at the same time). Further, looking at the aggregated AEIM matrix, Figure 4, for the time frame between 7am and 10pm (average over 200 simulation runs) this first intuition is strengthened.

From the A^{aai} matrix in Figure 4 we can observe that besides the obvious *Treatment*, *Assessment* and *Bed Placement* activities, *Wait In Bed* has a very high influence (in particular 4.5) on patients to be assigned to a room. Looking at the A^{aei} matrix it becomes apparent that rooms are the major cause for this (on average 571 time steps (minutes) per simulation run), although they appear not to be the main BN (451 time steps (minutes) in the *nbe* values vs. 521 aggregated across nurses and FT nurses). Hence, there are blocking effects in the model that lead to inactive periods of resources which possibly could be avoided. One may remember that the incentive for dispatching activities in the order *Bed Placement*, *Assessment* and *Treatment* was to ensure faster TTD times. However, it seems that this results in blocking of rooms and resources. Indeed, aiming to get patients out of rooms before assigning new patients to rooms, i.e. reversing the above order, results in shorter LOS and TTD times. In particular, the overall LOS can be reduced from 162.1 minutes to 144.4 minutes and the overall TTD from 116.1 minutes to 103.9 minutes.

The AEIM matrix that results from dispatching activities in the reverse order, i.e., *Treatment*, *Assessment*, *Bed Placement*, is illustrated in Figure 5. One can see that the influence of *Wait in Bed* on *Bed Placement* was reduced by approximately 60% and the main cause for unfulfilled requests shifts to nurses (including

Figure 3: AEIM matrices from 3pm-7pm for a single simulation run. Matrices A^{aai} and A^{eai} are flattened by reporting only relevant combinations.

FT nurses). AEIM enables the effect of BN behavior to be clearly identified when the reverse activity dispatch order is tested.

6.2 Analyzing the impacts of triage grades

It is a well known effect that high priority patients or tasks often block lower priority patients or tasks and cause longer LOS for the latter. Hence, priority accumulation techniques have gained more and more interest as approaches to overcome such effects (see for example Sengupta et al. (2011)). While priority accumulation is a suitable approach to deal with blocking based on priority, AEIM enables the detailed assessment of the impact of blocking between different patient categories. Expanding activity definitions to separate activities for each patient category enables measurement of such blocking behavior. Figure 6 shows the resulting AEIM matrix during day hours. It shows that in this particular setting high priority patients are not the major cause for longer waiting times of ESI3 and ESI4 patients. Although, as expected, they do have an influence, the major cause is the high volumes of ESI3 patients (self influence, a^{aai} value of 2.6), and, in particular, their blocking effect on ESI4 patients (a^{aai} value of 2.5).

	BedPlacement	Assessment	Treatment	FirstTreatmentES1	WaitInBed	BedPlacementFT	TreatmentFT	AssistFT	XRay	Activity Influenced	Activity Active	Doctors	Nurses	FTNurses	TriageNurses	AcuteRooms	MainRooms	FTRooms	XRayRooms
BedPlacement	4.7	5.3	9.1	0.6	4.5	0.0	0.0	0.0	4.4	28.6	458	0	391	526	0	0	571	0	0
Assessment	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	500	21	0	0	0	0	0	0	0
Treatment	0.6	0.7	1.2	0.1	0.0	0.0	0.0	0.0	0.0	2.7	672	475	354	478	0	0	0	0	0
FirstTreatmentES1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	67	5	7	9	0	4	9	0	0
WaitInBed	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	326	0	0	0	0	0	0	0	0
BedPlacementFT	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.1	0.0	0.4	156	0	0	207	0	0	0	0	0
TreatmentFT	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	173	0	0	0	0	0	0	0	0
AssistFT	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	355	0	0	0	0	0	0	0	0
XRay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2	494	0	0	0	0	0	0	0	161
Agg. Influence	5.3	6.0	10.4	0.7	4.5	0.1	0.2	0.1	4.6			235	521		343	11	451	18	160

Figure 4: Aggregated AEIM matrix from 7am to 10pm over 200 simulation runs (relevant elements only).

	BedPlacement	Assessment	Treatment	FirstTreatmentES1	WaitInBed	BedPlacementFT	TreatmentFT	AssistFT	XRay	Activity Influenced	Activity Active	Doctors	Nurses	FTNurses	TriageNurses	AcuteRooms	MainRooms	FTRooms	XRayRooms
BedPlacement	4.3	4.8	8.5	0.5	1.9	0.0	0.0	0.0	4.0	24.0	467	0	436	527	0	0	431	0	0
Assessment	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.2	509	62	0	0	0	0	0	0	0
Treatment	0.3	0.3	0.7	0.1	0.0	0.0	0.0	0.0	0.0	1.4	666	312	251	309	0	0	0	0	0
FirstTreatmentES1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	66.6	4	7	7	0	3	6	0	0
WaitInBed	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	188	0	0	0	0	0	0	0	0
BedPlacementFT	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.1	0.0	0.4	155	0	0	206	0	0	0	0	0
TreatmentFT	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	173	0	0	0	0	0	0	0	0
AssistFT	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	364	0	0	0	0	0	0	0	0
XRay	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.3	504	0	0	0	0	0	0	0	190
Agg. Influence	4.6	5.1	9.3	0.6	1.9	0.1	0.2	0.1	4.3			269	541		329	11	349	17	186

Figure 5: Aggregated AEIM matrix with reversed dispatching order from 7am to 10pm over 200 simulation runs (relevant elements only).

7 CONCLUSION AND FURTHER RESEARCH

In this paper we introduced the AEIM to analyze the behavior of complex systems and/or simulation models. The AEIM was developed as a consequence of existing dynamic bottleneck detection methods, in particular APM, being unsuitable as tools for revealing blocking effects within models that do not follow the strict resource/task couplings that are usually witnessed in line production. The main goal of the AEIM is to identify causes of blocking and starving effects and enable a deeper understanding of emerging behavior in models by providing simple methods to measure and visualize such effects. Hence, AEIM can be considered as an addition to classical KPIs, that are used to measure the performance of systems/models. Both the shortcomings of APM and the efficacy of AEIM are demonstrated on an example model of a

	BedPlacementESI1	BedPlacementESI2	BedPlacementESI3	BedPlacementESI4	AssessmentESI2	AssessmentESI3	AssessmentESI4	TreatmentESI1	TreatmentESI2	TreatmentESI3	TreatmentESI4	FirstTreatmentESI1	Activity Influenced
BedPlacementESI2	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.2	0.0	0.0	0.8
BedPlacementESI3	0.0	0.7	1.4	0.0	0.7	1.6	0.0	0.4	1.3	2.6	0.1	0.3	12.0
BedPlacementESI4	0.0	0.5	1.5	0.2	0.5	1.6	0.2	0.2	0.9	2.5	0.2	0.2	11.2
AssessmentESI2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
AssessmentESI3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1
AssessmentESI4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TreatmentESI1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
TreatmentESI2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.3
TreatmentESI3	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.1	0.3	0.0	0.1	0.9
TreatmentESI4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
FirstTreatmentESI1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Agg. Influence	0.0	1.3	3.1	0.2	1.4	3.5	0.2	0.7	2.4	5.8	0.4	0.6	

Figure 6: Aggregated AEIM matrix for all patient categories with reversed dispatching order from 7am to 10pm over 200 simulation runs.

typical ED. Results demonstrate the practicality of the AEIM for identifying dynamic bottlenecks caused by both the process sequence and prioritization. The use of the method within a real-world simulation study is beyond of the scope of this paper and left for further research.

REFERENCES

Abo-Hamad, W., and A. Arisha. 2013. "Simulation-Based Framework to Improve Patient Experience in an Emergency Department". *European Journal of Operational Research* 224(1):154–166.

Barrera-Diaz, C., M. Nakano, J. Oscarsson, S. Lidberg, and T. Sellgren. 2018. "Discrete Event Simulation Output Data-Handling System in an Automotive Manufacturing Plant". In *Procedia Manufacturing*, edited by M. Onori, L. Wang, X. V. Wang, and W. Ji, 23–30.

Chang, Q., J. Ni, P. Bandyopadhyay, S. Biller, and G. Xiao. 2007. "Supervisory Factory Control Based on Real-Time Production Feedback". *Journal of Manufacturing Science and Engineering* 129(3):653–660.

Currie, C. S. M., and R. C. H. Cheng. 2016. "A Practical Introduction to Analysis of Simulation Output Datat". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 118–132. Piscataway, New Jersey: Institute of Electrical and Electronics Engineering, Inc.

Day, T. E., A. R. Al-Roubaie, and E. J. Goldlust. 2013. "Decreased Length of Stay After Addition of Healthcare Provider in Emergency Department Triage: A Comparison Between Computer-Simulated and Real-World Interventions". *Emergency Medicine Journal* 30(2):134–138.

Furian, N., D. Neubacher, M. O'Sullivan, C. Walker, and C. Pizzera. 2018. "GEDMod Towards a Generic Toolkit for Emergency Department Modeling". *Simulation Modelling Practice and Theory* 87:239–273.

Furian, N., D. Neubacher, S. Vössner, M. O'Sullivan, and C. Walker. 2016. "HCDESLib - A Generic Open Source Toolkit For Health Care Simulation". In *The European Simulation and Modelling Conference 2016*, edited by J. Évora-Gómez and J. J. Hernández-Cabrera, 315–319. Ostend, Belgium: The European Simulation Society - The European Technology Institute.

Furian, N., M. O'Sullivan, C. Walker, S. Vössner, and D. Neubacher. 2015. "A Conceptual Modeling Framework for Discrete Event Simulation Using Hierarchical Control Structures". *Simulation Modelling Practice and Theory* 56:82–96.

Ghanes, K., O. Jouini, Z. Jemai, M. Wargon, R. Hellmann, V. Thomas, and G. Koole. 2014. "A Comprehensive Simulation Modeling of an Emergency Department: A Case Study for Simulation Optimization of Staffing Levels". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 1421–1432. Piscataway, New Jersey: Institute of Electrical and Electronics Engineering, Inc.

Gunal, M. M., and M. Pidd. 2006. "Understanding Accident and Emergency Department Performance Using Simulation". In *Proceedings of the 2006 Conference on Winter Simulation*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 446–452. Piscataway, New Jersey: Institute of Electrical and Electronics Engineering, Inc.

Heath, S. K., A. Buss, S. C. Brailsford, and C. M. Macal. 2011. "Cross-paradigm Simulation Modeling: Challenges and Successes". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. Creasey, J. Himmelspach, K. White, and M. Fu, 2788–2802. Piscataway, New Jersey: Institute of Electrical and Electronics Engineering, Inc.

Kuo, C.-T., J.-T. Lim, and S. M. Meerkov. 1996. "Bottlenecks in Serial Production Lines: A System-Theoretic Approach". *Mathematical Problems in Engineering* 2(3):233–276.

Law, A. M. 2015. "Statistical Analysis of Simulation Output Data: The Practical State of the Art". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 1810–1819. Piscataway, New Jersey: Institute of Electrical and Electronics Engineering, Inc.

Lim, M. E., A. Worster, R. Goeree, and J.-É. Tarride. 2013. "Simulating an Emergency Department: The Importance of Modeling The Interactions Between Physicians and Delegates in a Discrete Event Simulation". *BMC Medical Informatics and Decision Making* 13(1):59–70.

Parmenter, D. 2007. *Key Performance Indicators: Developing, Implementing, and Using Winning KPIs*. New York, USA: John Wiley & Sons, Inc.

Roser, C., and M. Nakano. 2015. "A Quantitative Comparison of Bottleneck Detection Methods in Manufacturing Systems with Particular Consideration for Shifting Bottlenecks". In *Advances in Production Management Systems: Innovative Production Management Towards Sustainable Growth*, edited by S. Umeda, M. Nakano, H. Mizuyama, H. Hibino, D. Kiritsis, and G. von Cieminski, 273–281. Cham, Switzerland: Springer International Publishing.

Roser, C., M. Nakano, and M. Tanaka. 2001. "A Practical Bottleneck Detection Method". In *Proceedings of the 2001 Winter Simulation Conference*, edited by B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, 949–953. Piscataway, New Jersey: Institute of Electrical and Electronics Engineering, Inc.

Roser, C., M. Nakano, and M. Tanaka. 2003. "Comparison of Bottleneck Detection Methods for AGV Systems". In *Proceedings of the 2003 Winter Simulation Conference*, edited by S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, 1192–1198. Piscataway, New Jersey: Institute of Electrical and Electronics Engineering, Inc.

Sengupta, S., M. Deneweth, and R. V. Til. 2011. "A Better Approach to Modeling Emergency Care Service". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. Creasey, J. Himmelspach, K. White, and M. Fu, 1202–1210. Piscataway, New Jersey: Institute of Electrical and Electronics Engineering, Inc.

Wang, Y., Q. Zhao, and D. Zheng. 2005. "Bottlenecks in Production Networks: An Overview". *Journal of Systems Science and Systems Engineering* 14(3):347–363.

AUTHOR BIOGRAPHIES

NIKOLAUS FURIAN is an Assistant Professor in the Department of Engineering and Business Informatics at TU Graz. He has a master degree in technical mathematics and a Phd in industrial engineering with focus on operations research. His main research interests are in the fields of simulation, optimization and data analytics. He can be contacted by email at nikolaus.furian@tugraz.at.

CAMERON WALKER is an Associate Professor in the Department of Engineering Science at the University of Auckland. He has Masters Degrees in Operations Research, Arts (Mathematics), Science (Statistics) and a PhD in Mathematics. His main research interest is improving health delivery processes. He can be contacted by email at cameron.walker@auckland.ac.nz.

MICHAEL O'SULLIVAN is a Senior Lecturer in the Department of Engineering Science at the University of Auckland where he completed a degree in Mathematics and Computer Science and a Masters in OR. Afterwards, he obtained a MS in Engineering-Economic Systems and OR, and a PhD in Management Science and Engineering at Stanford University in the US. Upon his return to New Zealand, he joined the Department of Engineering Science where he co-founded the research group Operations Research Union Analytics (ORUA). He can be contacted by email at michael.osullivan@auckland.ac.nz.

DIETMAR NEUBACHER is a Senior Researcher at the Department of Engineering and Business Informatics at the TU Graz. He has a master degree in mechanical engineering and business economics and a Phd in industrial engineering. His main research focus is on modeling and simulation of socio-technical systems and can be contacted by email at dietmar.neubacher@tugraz.at.

CLEMENS GUTSCHI is a research assistant at the Department of Engineering and Business Informatics at the TU Graz. He has a masters degree in mechanical engineering and business economics and attends the doctoral school in techno-economics with focus on the maintenance strategy selection problem. He can be contacted by email at clemens.gutschi@tugraz.at.