

Grzegorz Dobrowolski*, Edward Nawarecki*

Sytuacje kryzysowe w systemach agentowych

1. Wstęp

Prowadzone w pracy rozważania odnoszą się do klasy inteligentnych systemów zdecentralizowanych, spełniających paradygmaty agentowości. Mogą to być zarówno systemy projektowane jako agentowe (działające w świecie wirtualnym, np. system usług informacyjnych w sieci komputerowej, przedsiębiorstwo wirtualne), jak również systemy działające w świecie rzeczywistym (usługi transportowe, złożone systemy produkcyjne itp.).

W systemach takich (zarówno wirtualnych, jak i rzeczywistych) występować mogą sytuacje o charakterze kryzysowym, wywołane czynnikami zewnętrznymi (niepożądana ingerencja, siły przyrody) lub wewnątrzsystemowymi (deficyt zasobów, uszkodzenia lokalne). W przyjętym rozumieniu, powstanie kryzysu interpretowane jest jako zagrożenie utratą (częściową lub całkowitą) funkcjonalności systemu. Trudności z rozpoznaniem kryzysu, oceną jego skutków lub podjęciem działań zapobiegawczych (antykryzysowych) wynikają z własności systemu (autonomia decyzji agentów, brak informacji globalnych) oraz jego dynamiki (skutki mogą pojawić się dopiero po pewnym okresie działania systemu) [4, 1].

Zamierzeniem pracy jest zaproponowanie formalizacji opisu agenta i systemu agentowego, która może posłużyć do analizy działania systemu, ze szczególnym uwzględnieniem sytuacji kryzysowych. Proponowane podejście umożliwia sprecyzowanie zadań systemu monitorującego działania systemu podmiotowego, a następnie stworzenie modelu symulującego [3, 5] jego działania w określonych sytuacjach kryzysowych. Wynikiem symulacji są pewne scenariusze dynamicznego rozwoju takich sytuacji. Analiza uzyskanych scenariuszy pozwala na poszukiwanie strategii przeciwdziałania kryzysom lub przynajmniej ograniczających ich skutki.

Rozważania prowadzone są na takim poziomie ogólności, aby możliwe było ich dostosowanie do specyfiki konkretnego zastosowania.

2. Modele agenta i systemu

Dalej przedstawiony zostanie model agenta będący konsekwencją *wyjęcia* odpowiedniego systemu z otoczenia i wyspecyfikowania obserwowanych oddziaływań. W ten sposób

* Katedra Informatyki, Akademia Górniczo-Hutnicza w Krakowie

model ujmuje przede wszystkim zdolność agenta do współistnienia z innymi agentami, czyli te cechy, które stanowią o sposobie wiązania się agentów w systemy, ale również w ogólnym zakresie definiuje jego (abstrakcyjną) architekturę wewnętrzną. Pełny opis modeli wprowadzonych poniżej znaleźć można w [2].

Definicja 1. *Agentem Λ nazywamy trójkę*

$$\Lambda = \{A, S, F \subset S \times A \times S\} \quad (1)$$

gdzie:

- A – skończony zbiór akcji (działań elementarnych) agenta;
- S – skończony zbiór stanów wewnętrznych agenta;
- F – trójczłonowa relacja określająca możliwe następstwa stanów i akcji agenta, interpretowana w następujący sposób: w określonym stanie agent może wykonać akcję (drugi człon relacji), która odpowiednio zaprowadzi go do nowego stanu (trzeci człon relacji).

Zachowanie związków przyczynowo-skutkowych wymaga, aby F była taka, że zachodzi

$$(s, a, s_1) \in F \wedge (s, a, s_2) \in F \Rightarrow s_1 = s_2 \quad (2)$$

Relacja F , wiążąca akcje (niepodzielne) ze stanami, wskazuje nie tylko, do jakiego stanu agenta prowadzi wykonanie którejś z nich, ale również określa, jaki podzbiór akcji jest w danym stanie wykonalny (dopuszczalny), czyli możliwości wyboru, przed jakim staje agent. Spoiwem systemu mogą być tylko ich akcje.

Definicja 2. *Systemem agentowym Ω nazywamy dwójkę*

$$\Omega = \{\{\Lambda^i\}_{i=1, \dots, N}, I\} \quad (3)$$

którą stanowią skończony zbiór agentów systemu $\{\Lambda^i\}_{i=1, \dots, N}$ oraz relacja zwana relacją współdziałania, o postaci

$$I = \bigcup_{\substack{i, j=1, \dots, N \\ i \neq j}} I^{ij} \quad (4)$$

przy czym: $A^i \times A^j \supset I^{ij} \ni (a^i, a^j) : A^i \in \Lambda^i; A^j \in \Lambda^j; \Lambda^i, \Lambda^j \in \Omega$.

Relacja I jest symetryczna: $(a^i, a^j) \in I \Rightarrow (a^j, a^i) \in I$.

Relacja współdziałania określa potencjalne możliwości współdziałania agentów w obrębie systemu. Faktyczne współdziałanie następuje, gdy znajdujące się w relacji I akcje są wykonywane przez agentów, przy czym – co oczywiste – żadna z akcji składowych nie może być wykonana niezależnie. W konsekwencji, działanie systemu może być zdefiniowane jako złożenie działań agentów wchodzących w jego skład.

3. Cele i decyzje w systemie agentowym

Wychodząc od założonej użyteczności systemów agentowych, jak również ich istoty, stwierdzić można, że z istnieniem każdego z takich systemów związane są:

- cel globalny systemu oraz ewentualnie związane z nim cele poboczne, czyli powody, dla których system został zbudowany;
- cele indywidualne agentów, które każdy z nich posiada z racji założonej autonomii i które mogą być *rozbieżne* (zakłada się poszukiwanie i realizację kompromisu) pomiędzy sobą, jak również względem celu globalnego systemu.

Cele powyższe odpowiednio przekładają się na strategie systemu i jego agentów, które zastosowane, zapewnić powinny osiągnięcie tych celów.

Realizacja celu, strategii indywidualnej, czyli decyzje o wykonaniu określonych akcji są efektem **mechanizmu wyboru akcji** agenta, który formalnie wyraża się następująco.

Definicja 3. *Agentem ze strategią Λ_U nazywamy rozszerzenie agregatu Λ agenta z definicji 1, mające postać*

$$\Lambda_U = \{\Lambda, \{u_s\}_{s \in S}\} \quad (5)$$

w którym zbiór $\{u_s\}_{s \in S}$, nazywany strategią agenta, jest rodziną odwzorowań – funkcji wyboru takich, że

$$\forall s \in S \exists u_s : A \supset A_s \rightarrow a \in A_s; A, S \in \Lambda \quad (6)$$

Opis celu indywidualnego oraz procesu podejmowania decyzji przez agenta w przypadku ogólnym poprowadzony może być głębiej jedynie pod warunkiem zdefiniowania stanu agenta oraz pozostałych elementów z definicji 3. I tak przykładowo, jeżeli jako komponent stanu przyjmie się zasób informacyjny (wiedzę) agenta, która będzie się zmieniać w trakcie jego działania, zmieniać się będą również obserwowane z zewnątrz wybory agenta – agent będzie „zmieniał poglądy”. Jeżeli stan agenta będzie można ocenić liczbowo, funkcję wyboru będzie można definiować na bazie teorii (funkcji) użyteczności.

Jednocześnie zwrócić trzeba uwagę na fakt, że opis celu globalnego i strategii całego systemu również musi abstrahować od reprezentacji stanu i realizacji mechanizmu wyboru akcji agentów wchodzących w skład systemu. Musimy zatem posłużyć się tutaj pojęciem *działania systemu agentowego*.

Definicja 4. *Strategią (globalną) systemu agentowego Ω nazywamy funkcję wyboru u_Ω postaci*

$$u_\Omega : D \rightarrow P(D) \quad (7)$$

gdzie D jest zbiorem wszystkich możliwych działań systemu ($P(D)$ – rodzina podzbiorów D). Jednocześnie odpowiedni (wybrany) podzbiór D nazywamy celem globalnym.

Jakkolwiek nie zostało to zastrzeżone w definicji, funkcja u_Ω powodów interpretacyjnych powinna dawać w wyniku niepuste i niezbyt liczne podzbiory D .

Często strategia definiowana jest w postaci postulatu w stosunku do zadanej obliczalnej charakterystyki działania systemu. Dwie z możliwości podać można od razu. Pierwsza to postulat osiągalności pewnego stanu końcowego lub konstrukcji działania systemu prowadzącego do tego stanu formułowany również jako warunek integralności systemu. Druga to możliwość formułowania szczegółowych postulatów w oparciu o bilanse zasobów w systemie. Mają one kształt zrelaksowanych zadań optymalizacji, pożądane zaś działanie systemu ma „lokować się w pobliżu optimum”.

Rozważmy możliwości implementacji obu rodzajów strategii. W przeciwieństwie do strategii globalnej, strategie lokalne albo stanowią składowe agentów, albo dają się w nie prosto wbudowywać. Na mocy założenia o racjonalności agentów przyjmuje się, że następnie strategie te będą realizowane.

Wobec autonomii agentów, już precyzyjne określenie strategii globalnej wydaje się trudne (nie jest konstruktywna również definicja 4), zaś jej realizacja (zorganizowanie współdziałania agentów) jest faktycznie podstawowym problemem konstrukcji systemów agentowych.

4. Aspekty modelowe sytuacji kryzysowych

Podstawą dla dalszych rozważań będzie następująca ogólna charakterystyka sytuacji kryzysowych w odniesieniu do systemów agentowych.

Sytuacja kryzysowa to stan lub ciąg stanów systemu, które naruszają lub prowadzić mogą do naruszenia realizacji celów systemu zarówno globalnego, jak i celów lokalnych poszczególnych agentów.

Sytuacje kryzysowe mogą być zatem:

- lokalne (dotyczące jednego agenta),
- globalne (w tym również grupy agentów).

Ujawnienie się lokalnej sytuacji kryzysowej może w perspektywie czasowej pociągnąć za sobą kryzys całego systemu, ale często możliwości funkcjonalne agentów wystarczają do oddalenia kryzysu lokalnego. Zjawisko to wynika wprost z podstawowych właściwości systemu. Można powiedzieć, że jest on już wyposażony w mechanizmy antykryzysowe, ale o ograniczonym zakresie stosowania. Groźba kryzysu globalnego natomiast zwykle wymaga opracowania specjalnych mechanizmów systemowych.

Zaliczenie sytuacji kryzysowej wynikłej w grupie agentów (podsystemie) do globalnych wynika z podobnych możliwości określenia stanu i tego, że kryzys w podsystemie może być obserwowany w konsekwencji celu, którym jest zdekomponowany cel globalny lub jeden z pobocznych celów całego systemu agentowego.

Powyższa ogólna charakterystyka sytuacji kryzysowych określa następujące warunki zarządzania nimi:

- istnienie możliwości obserwacji (monitorowania) stanu systemu poprzez obserwację stanów poszczególnych agentów;
- przyjęcie odpowiedniego sposobu ewaluacji stanu dla uzyskania operacyjnego kryterium zaistnienia sytuacji kryzysowej;
- istnienie środków pozwalających na wyprowadzenie systemu z sytuacji kryzysowej.

Stopień realizacji powyższych warunków może być wyznacznikiem odporności systemu na sytuacje kryzysowe. Zwykle zachodzi, że *podstawowa* konstrukcja systemu już zawiera elementy mechanizmów antykryzysowych. Mieszczą się one w konstrukcji pojedynczego agenta, wymianie informacji pomiędzy agentami, czy organizacji systemu (podsystemów).

Przedyskutujmy te warunki najpierw dla przypadku sytuacji kryzysowych lokalnych, a następnie globalnych.

W oczywisty sposób **agent** monitoruje swój stan, jak również wyposażony jest w sposoby jego ewaluacji. Po pierwsze, dokonuje w stanie s określenia zbioru A_s , którego istotna redukcja może być sygnałem sytuacji kryzysowej. Jeżeli w repertuarze agenta istnieją akcje specjalne w rodzaju: *nic nie rób*, samodestrukcji, zawężenie wyboru do którejś z nich, to nie tylko kryzys, ale również odpowiednie postępowanie wobec niego.

Po drugie, strategia agenta u_s w tym stanie to również ewaluacja tego stanu, tyle że zorientowana na określenie działania. I tak, jeżeli w ramach obliczania u_s określany jest ranking akcji do wykonania z wagami (użyteczności), wówczas odpowiednie wartości mogą być zastosowane jako kryterium sytuacji kryzysowej (spadek użyteczności działania). Wreszcie, jeżeli oba powyższe sposoby są niewystarczające, do funkcji wyboru można wbudować dodatkową składową, która faktycznie będzie monitorować stany krytyczne i ewentualnie powodować wybranie którejś z dodanych akcji antykryzysowych.

Podobna analiza w odniesieniu do globalnych sytuacji kryzysowych okazuje się trudniejsza, co ma swoje źródło w problemie *uchwycenia* stanu **systemu agentowego**. Stan systemu może być bowiem zdefiniowany jako złożenie stanów agentów, jednakże precyzyjne operacyjne określenie tego stanu na ogół nie jest możliwe, co ma swoje źródło w specyficznych właściwościach systemu:

- brak mocnego mechanizmu synchronizacji – trudności w ustaleniu jednoczesności stanów poszczególnych agentów;
- wysoka wymiarowość stanu systemu – wysokie koszty akwizycji informacji o stanie;
- autonomia agentów – agent ujawnia swój stan w stopniu minimalnym, zwykle jedynie tyle, aby zadość uczynić zasadom działania systemu.

W ogólnym przypadku założyć można, że agent j ujawnia jedynie pewną podprzestrzeń swojego stanu s^{*j} lub pewną jego ewaluację $v^j(s^j)$, przy czym interpretacja tych informacji jest znana w systemie. Uznaje się przy tym, że jakkolwiek projekcja stanu s^{*j} jest *czystym* raportem, ewaluacja może być obciążona pewnym *subiektywizmem*. Zwrócić należy również uwagę na to, że w stanie agenta zmagazynowana jest informacja o historii jego działania, zatem odpowiednia ewaluacja nosi cechy dynamicznej.

Połączenie opisów stanów takiej postaci, dostarczonych przez agentów i uznanych ponadto za jednoczesne, przyjęte być musi za opis stanu systemu. W zakresie architektury systemu uzyskanie takiego opisu wymaga zaprojektowania i realizacji specjalnego podsystemu monitorującego – agenta (agentów) zbierającego informacje z wykorzystaniem specjalizowanych protokołów interakcyjnych.

Założmy w dalszym ciągu, że podsystem monitorujący istnieje i możliwa jest następująca, specyficzna ewaluacja stanu systemu agentowego

$$v_{\Omega} = v_{\Omega}(s^{*j}, \dots, v^j(s^j), \dots) : A^j \in \Omega \quad (8)$$

Ewaluacja ta, podobnie jak ewaluacje poszczególnych agentów, może być zorientowana (zależnie od założeń w stosunku do podsystemu monitorowania) na sytuacje kryzysowe. Nadanie funkcjom ewaluacji odpowiedniej postaci oraz odpowiednia definicja zbiorów ich wartości umożliwi ukierunkowane śledzenie stanów sytemu.

Przyporządkowując funkcjom wartości lingwistyczne, można na przykład uzyskać charakterystykę stanów jako:

- „normalne”,
- „preferowane”,
- „zagrożenia”,
- „kryzysu”.

Śledzenie może tu być również rozumiane jako zapamiętywanie danych zebranych i wprowadzenie ich do funkcji ewaluacji. Wymaga to oczywistych zmian we wzorze (8).

Podążając za możliwościami formułowania celu globalnego można wyróżnić następujące grupy sytuacji kryzysowych – bezpośrednich i pośrednich zagrożeń dla funkcjonalności systemu, wyrażanych odpowiednio przez niedostępność akcji określonych agentów (lub ich samych) oraz naruszenia bilansu zasobowego, które następnie przełożyć się mogą na deficyt funkcjonalny.

Wystąpienie obu tych przypadków może być wykryte przez system monitorowania czy to przez śledzenie ewaluacji indywidualnych wskazujących na utratę funkcjonalności, czy dystrybucji pewnego zasobu krytycznego dla pracy agenta lub systemu.

5. Struktura systemu zarządzania kryzysami

Przedstawiony opis działania agenta oraz całego systemu agentowego stanowi punkt wyjścia do zaproponowania rozwiązań w zakresie struktury i funkcjonalności systemu zarządzania w sytuacjach kryzysowych. Możliwość wystąpienia sytuacji kryzysowych może być bardzo wiele – zależy to zarówno od przeznaczenia systemu agentowego, jak i zastosowanych w nim rozwiązań. Dlatego też trudne (a może nawet niecelowe) wydaje się poszukiwanie uniwersalnej koncepcji zarządzania w sytuacjach kryzysowych. W dalszym ciągu przedstawiono dwa charakterystyczne przypadki, uzupełnione kilku pomysłami na poziomie implementacyjnym, które mogą być zastosowane w sposób wariantowy.

Z rozważań prowadzonych w punktach poprzednich wynika, że problem zarządzania sytuacjami kryzysowymi zwiiera w sobie dwa istotne dające się rozdzielić podproblemy:

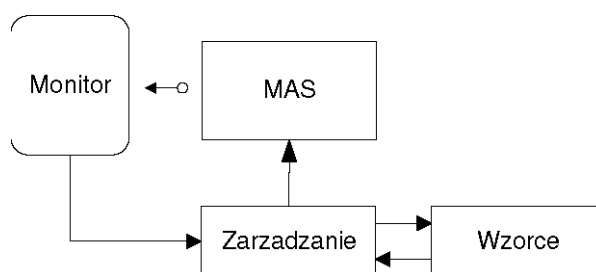
- 1) monitoringu stanu systemu,
- 2) zarządzania systemem.

W odniesieniu do przypadku, w którym agenci działają w środowisku wirtualnym (systemie komputerowym), problem pozyskania danych, które w formie zagregowanej reprezentują stan systemu agentowego jest zagadnieniem standardowym informatyki; istotna może być jedynie liczba danych, co wpływa na przyjęte rozwiązania techniczne. Proble-

mem jest natomiast algorytmiczne rozwiązanie kwestii rozpoznawania sytuacji kryzysowych (lub sygnalizacji zagrożenia powstania takich sytuacji) i reakcji na nie. Działanie odpowiedniego bloku zarządzania może przebiegać według różnych zasad, wśród których można wskazać:

- zarządzanie w oparciu o wcześniej opracowane wzorce sytuacji kryzysowych,
- zarządzanie *on-line* poprzez kreowanie odpowiednich reakcji (zmian) w systemie agentowym.

Ogólny schemat zarządzania dla omawianego przypadku przedstawiono na rysunku 1.



Rys. 1. Struktura zarządzania dla przypadku systemu wirtualnego
Objaśnienia w tekście

System agentowy (MAS) jest obserwowany przez podsystem monitoringu (Monit), z którego korzysta moduł zarządzania, wykorzystujący wzorce. Rozwiązanie jest właściwe, m.in. dla przypadków, gdy system agentowy przeznaczony jest do pełnienia funkcji obliczeniowych (*soft-computing*) lub roli modelu symulującego w trybie *off-line* działanie pewnego systemu rzeczywistego (np. szeregowanie zadań lub sieć transportową w planowaniu).

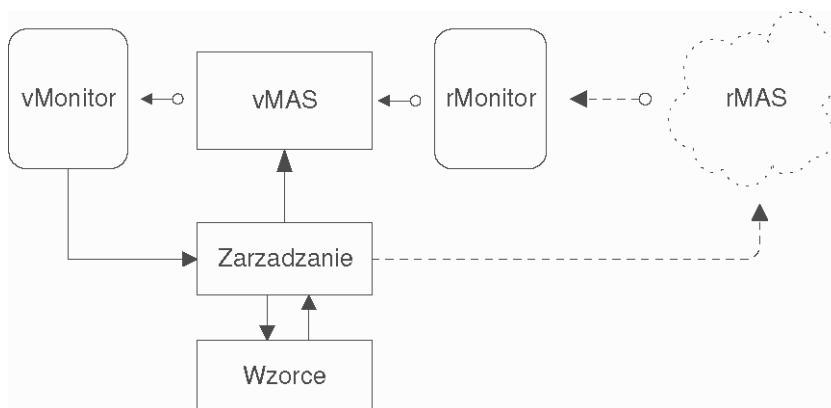
Problem monitoringu staje się bardziej skomplikowany, gdy agenci w obserwowanym systemie posiadają silną autonomię, czyli mogą nie poddawać się bezpośredniej obserwacji lub obserwacja ta jest znacznie utrudniona. Sytuacja taka może występować m.in. w otwartych i heterogenicznych systemach informacyjnych (dostęp do informacji może być utrudniony chociażby ze względu na brak wspólnej ontologii), lub w systemach rzeczywistych traktowanych w trybie *on-line* (bezpośrednia obserwacja ruchu w sieci komunikacyjnej, działanie grupy zakładów produkcyjnych lub przedsiębiorstw handlowych, monitorowanie środowiska naturalnego).

W warunkach trudności w odtworzeniu stanów agentów oraz ograniczonych możliwości ingerencji w ich działania uzasadnione staje się rozbudowanie schematu z rysunku 1 o dodatkowy blok, realizujący ładowanie *on-line* danymi systemu agentowego, który funkcjonuje jako agentowy model rzeczywistości (patrz rys. 2).

Mamy zatem dwa systemy monitorujące:

- 1) rMonitor,
- 2) vMonitor.

Pierwszy z nich służy do pozyskiwania informacji dostępnych w systemie rzeczywistym (rMAS), drugi zaś do agregacji danych generowanych w modelu (vMAS). Wzbogaceniu podlegają również funkcje realizowane w bloku zarządzania – oprócz zarządzania modelem (vMAS), mającym na celu badanie różnych wariantów działań antykrzysowych, pojawia się egzekwowanie (inicjowanie egzekwowania) w warunkach rzeczywistych (rMAS) strategii wypracowanych w modelu.



Rys. 2. Struktura zarządzania dla przypadku systemu rzeczywistego
Objaśnienia w tekście

Warto zwrócić uwagę, że struktury przedstawione na rysunkach 1 i 2 ilustrują funkcjonalności występujące w prezentowanym podejściu, natomiast w konkretnej realizacji bloki te nie muszą być wyodrębnione. W szczególności system monitoringu (vMonit) może być pewną „nakładką” na monitorowany system agentowy. Przy takim rozwiązaniu system monitorowany może być interpretowany jako środowisko systemu monitorującego.

6. Podsumowanie

Prowadzone w artykule rozważania dotyczyły zastosowania podejścia agentowego do problemu zarządzania w sytuacjach kryzysowych.

W proponowanym podejściu wyróżnić można dwa aspekty:

- 1) postępowanie w sytuacjach zagrożenia utratą funkcjonalności samego systemu agentowego, który z reguły posiadać będzie charakter wirtualny (np. systemy informacyjne lub obliczeniowe realizowane przy zastosowaniu technologii agentowych);
- 2) wykorzystanie systemu agentowego do podejmowania (wspomagania) decyzji w sytuacjach kryzysowych występujących (lub mogących zaistnieć) w świecie realnym (m.in. w złożonych kompleksach działań, systemie komunikacyjnym, energetycznym lub środowisku ekologicznym).

Jakkolwiek tryb postępowania w każdym z tych przypadków będzie inny, jednakże środki techniczne wykorzystywane do zarządzania będą w znacznej mierze podobne, co pozwala na uogólnienie przynajmniej części rozważań.

Prowadzone aktualnie w Katedrze Informatyki AGH badania dotyczą realizacji poszczególnych bloków składających się na systemy z rysunków 1 i 2. W szczególności uruchamiany jest agentowy symulator ciągów technologicznych oraz sieci usług transportowych oraz pilotowe wersje systemów monitorujących. Równolegle prowadzone są prace dotyczące algorytmów zarządzania w sytuacjach kryzysowych.

Literatura

- [1] Collins J., Tsvetovas M., Sundareswara R., van Tonder J., Gini M., Mobasher B.: *Evaluating risk: flexibility and feasibility in multi-agent contracting*. In: O. Etzioni, J.-P. Miiller, J.M. Bradshaw, (Eds), Proceedings of the Third International Conference on Autonomous Agents (Agents'99), Seattle, WA, USA, 1999. ACM Press, 350–351
- [2] Dobrowolski G.: *Technologie agentowe w zdecentralizowanych systemach informacyjno-decyzyjnych*. Wyd. 1. Kraków, UWND AGH 2002
- [3] Uhrmacher A.M., Gugler K.: *Distributed, parallel simulation of multiple, deliberative agents*. In: D. Bruce, L. Donatiello, S. Turner, (Eds), Proceedings of the 14th Workshop on Parallel and Distributed Simulation (PADSOO), Los Alamitos, CA, 2000, IEEE Press, 101–110
- [4] Wu S., Soo V.: *Risk control in multi-agent coordination by negotiation with a trusted third party*. In: D. Thomas (Ed.), Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI-99-Voll), Morgan Kaufmann Publishers, 1999, 500–505
- [5] Zhao Z., Belleman R.G., van Albada G.D., Sloot P.M. A.: *Scenario switches and state updates in an agent-based solution to constructing interactive simulation systems*. In: Proceedings of the Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS2002), 2002, 3–10

