TRANSNAV
http://www.transnav.eu
the International Journal
on Marine Navigation
and Safety of Sea Transportation
Volume 7
Number 2
June 2013

DOI: 10.12716/1001.07.02.18

Evolutionary Ship Track Planning within Traffic Separation Schemes – Evaluation of Individuals

R. Szlapczynski
Gdansk University of Technology, Poland

ABSTRACT: The paper presents an extended version of the author's Evolutionary Sets of Safe Ship Trajectories method. The method plans safe tracks of all ships involved in an encounter including speed reduction maneuvers, if necessary, and taking into account Rule 10 of COLREGS, which specifies ships' behavior within Traffic Separation Schemes governed by IMO. The paper focuses on the evaluation phase of the evolutionary process and shows how fitness function is designed to compare various possible tracks as well as to assess the quality of a final solution. The impact of the fitness function on the method's results is illustrated by examples.

1 INTRODUCTION

In earlier papers the author has presented the Evolutionary Sets of Safe Ship Trajectories (ESoSST) method (Szlapczynski 2011, Szlapczynski & Szlapczynska 2011, Szlapczynski & Szlapczynska 2012). This method, instead of finding the optimal own track for the unchanged courses and speeds of other ships, was searching for an optimal set of safe trajectories of all ships involved in an encounter, combining evolutionary approach (Smierzchalski & Michalewicz 2000) with goals typical for methods based on games theory (Lisowski 2007). The method has been further developed in (Szlapczynski 2012) to support Traffic Separation Schemes (TSS) (Anwar & Khalique 2006) and thus to be applied in Vessel Traffic Service (VTS) centers.

The method uses evolutionary algorithm which works as follows. First, the initial population of individuals (each being a potential solution to the problem) is generated. It includes individuals consisting of tracks being straight segments and tracks generated randomly as well as automatically determined TSS-compliant tracks. Each track is a sequence of nodes containing geographical coordinates. The initial population is a subject to subsequent iterations of evolutionary algorithm. Each of these iterations consists of the following steps:

- Evolutionary operations: individuals (sets of tracks) are modified by means of random mutation operators as well as specialized operators dedicated to the problem. The latter include collision-avoidance operators and TSS-compliance operators.
- Reproduction: pairs of parents are selected from all of the individuals and offspring is produced as a result of their mating. The offspring inherits some features from each parent. Three types of crossover operators are used here: random track exchange, one-point track crossover and intermediate recombination of nodes.
- Evaluation: each of the individuals (including parents and the offspring) is assigned a value of a fitness function, which reflects the quality of the solution represented by this individual. This involves detecting and penalizing all collisions as

well as penalizing violations of COLREGS Rules 13 to 17 and TSS violations (Rule 10 of COLREGS).
- Succession: the next generation of individuals is selected. The selection is based on the results of the evaluation. The individuals are chosen randomly, with the probability strictly depending on the fitness function value.

Normally, the evolutionary algorithm ends when one of the following happens:
- maximum acceptable time or number of iterations is reached,
- the satisfactorily high value of fitness function has been reached by one of the individuals,
- the evolution ceases to bring any improvement.

Recently the method has been further extended, the main new feature being planning speed reduction maneuvers within TSS, when necessary. The extensions have resulted in changes in the evaluation phase of the evolutionary process. The fitness function had to be redesigned to address these new issues. The paper shows how the new fitness function is able to compare various possible ship tracks as well as to assess the quality of a final solution. The paper is organized as follows. First, the optimisation problem is presented (Section 2). Then it is explained why the evaluation phase is crucial to the process (Section 3). This is followed by a discussion of the evaluation problems, which have to be solved (Section 4). The construction of the fitness function is shown next (Section 5) and some example results of applying this fitness function are provided (Section 6). Finally a summary and conclusions are given.

2 OPTIMIZATION PROBLEM

It is assumed that we are operating in a good visibility when COLREGS rules 13-17 abide and within a TSS governed by IMO, where Rule 10 additionally abides. The goal is to find a set of tracks, which minimizes the average time loss or way loss spent on manoeuvring, while fulfilling the following conditions:
- none of the stationary constraints (including TSS Inshore Traffic Zone [ITZ] and separation zones) are violated,
- none of the ship domains (Coldwell 1983) are violated,
- the course alteration should not be too small or too large (minimum and maximum alteration values are configurable and by default are set to 15 and 60 degrees respectively),
- a ship only manoeuvres when she is obliged to and, in case of head-on and crossing encounters, manoeuvres to starboard are favoured over manoeuvres to port,
- COLREGS rules (Cockcroft & Lameijer 2011, COLREGS 1972) are not violated (especially Rule 10 and Rules 13 to 17),
- speed alterations are not to be applied unless necessary (collision cannot be avoided by a configured maximum course alteration value),
- if speed alteration has to be applied, the number of speed alterations should be minimized (e.g. a ship can reduce her speed to avoid collision and get back to a normal speed once the situation is safe again).

It is also assumed that we are given the following data:
- stationary constraints (landmasses and other obstacles and the locations and parameters of each TSS's parts),
- positions, courses and speeds of all ships involved,
- simplified data on ships' manoeuvrability.

Some of the above mentioned parameters are provided either by ECDIS (landmasses and TSS coordinates) or by Automatic Identification System (AIS); motion parameters also by Automatic Radar Plotting Aid (ARPA). At present, no information on ship's manoeuvring abilities is provided in AIS messages, however these data can be directly obtained by a VTS operator from a ship's navigator and perhaps in future the content of AIS messages will be extended to include more information.

3 WHY IS EVALUATION PHASE CRUCIAL?

In general evaluation is necessary for converging to a solution and for telling how good this solution is according to the given set of criteria and constraints (Michalewicz & Fogel 2004). The most important purposes of evaluation are as follows.
1 Comparing between various individuals and making sure that a progress is made that way, because otherwise the evolutionary process will not converge at all.
2 Guarantying convergence to a "good" optimal solution. Otherwise we might face a situation when we converge to a solution that is optimal according to the given fitness function but turns out to be unacceptable, because our fitness function does not reflect our set of optimisation criteria well enough.
3 Deciding when our best individual is close enough to the optimal solution, so that we can return a solution to the user faster and save on computational time.
4 Deciding when our solution (best set of ship tracks) is acceptable and when it is not. Knowing that a solution is unacceptable we might apply a speed reduction (if possible) or inform that finding the demanded solution is not possible for the given data (e.g. because of ship domains being too large when compared with the given width of a lane or with the current distances between ships).

While the first three tasks are typical for evolutionary processes, the fourth one is more challenging. Solving it is absolutely necessary for making a decision on speed reduction. If there are ship domain violations in a solution, than we know for certain that this solution is incorrect and speed reduction should probably be applied. Unfortunately, usually there will be no ship domain violations in the solution, because the method will rather find a safe track at the cost of large way loss. Therefore the decision on the speed reduction must be made on the basis of fitness function value alone, rather than on the basis of the accompanying data on registered collisions. In practice, it is the unacceptably low total fitness function value that will always trigger an attempt to improve the solution by reducing the speed of the ship that has the lowest fitness value assigned to its track.

4 EVALUATION PROBLEMS AND THEIR SOLUTIONS

To make sure that fitness function satisfies the needs specified in the previous section, the following problems must be addressed:

1. How to compare unacceptable individuals with each other (differ between various levels of unacceptable individuals and decide which one is ‘less unacceptable’ and which one is ‘more unacceptable’)?
2. How to compare tracks that use traffic lanes with ones, which avoid lanes completely?
3. How to compare tracks which involve speed reduction with the ones that do not?
4. How to decide, that we are close enough the optimum? If we want a normalized fitness function than value ‘1’ should be assigned to an ideal solution, but we do not know what the ideal solution is, otherwise we wouldn't have to search for it.
5. How to assign fitness function values in such a way that values below certain threshold would be unacceptable for certain?

The answer to the first problem has been introducing a number of diagnostic factors: static constraint factor, collision avoidance factor, COLREGS-compliance factor and TSS-compliance factor, all of which reflect various conditions that have to be met. The first three factors have already been described in the author’s earlier papers. In general, each of these four factors has been assigned a different degree of penalty for condition violations:

- static constraint violations (penalized most severely because they have to be avoided at all cost),
- collisions with other ships (penalized slightly less severely because they might sometimes be eliminated as a side effect of avoiding static constraint violations),
- violations of Rules 13-17 of COLREGS (penalized moderately, because they are secondary when compared to collisions),
- violations of Rule 10 of COLREGS (penalized depending on a particular class of violation – e.g. moving against the traffic direction is penalized nearly as severely as violating static constraints).

The second of the above listed issues has been solved by introducing a lane encouragement factor (a component of the TSS-compliance factor), which is used for encouraging the method to plan trajectories, which use traffic lanes. For each track a percentage of the track’s length that transits through a traffic lane is determined and used for estimation of the track’s quality.

As for the third issue, the tracks, which utilize speed reduction and those that do not, are not compared with each other at all by the ESoSST method. The method tries to find safe tracks without speed reduction first and only if it fails to do so, speed reduction is applied. In such cases there is no penalty for speed reduction, because speed reduction is then treated as a necessity.

The fourth issue is strictly connected to the normalization of fitness function. In earlier versions of the ESoSST method track fitness was relatively easy to normalise. A track economy factor and various compliance factors were all from the <0,1> range and the track fitness function was simply a product of them all. However, once the TSS-compliance factor has been introduced, the normalization is more complex, because of the lane encouragement factor. If we penalize trajectories for not using traffic lanes, we might end up with some safe tracks being assigned very low track fitness values. On the other hand, if we reward using traffic lanes, some of the tracks may have their fitness values larger than 1, or close to 1 despite some obvious faults. Therefore a concept of reference track fitness has been introduced. Although we do not know the ideal track, we can try to determine the upper boundary of its fitness. This upper boundary must be carefully placed: placing it too high results in underestimating future solution, placing it too low – in overestimating the solution. Therefore the upper boundary – a reference track fitness value – is determined as follows. An optimal track of a particular ship is sought for, totally ignoring all other ships and collision avoidance rules. Due to ignoring other ships, we avoid way loss that is usually an effect of collision avoidance and we obtain the shortest track, which meets static constraints and is TSS-related constraints. If we divide a fitness value of any track by this track’s reference fitness value, we will get the desired normalized fitness value. This normalized fitness value approaches 1 as the track gets closer to the reference track (the one with no way loss form collision avoidance). Detailed formulas for normalized fitness function are provided in the next section.

The last of the listed problems – deciding when a track is unacceptable – is also connected to normalization and partly solved by the reference fitness value. When reference fitness value is used for normalizing a track’s fitness, we know that the difference between current fitness value and value ‘1’ can only be blamed on way loss due to collision avoidance. Therefore it is merely a question of settings (personal choice) how much way loss can be accepted before speed reduction is applied. In the method is has been assumed, that by default, a 5% way loss (fitness value of 0.95) can be accepted and any larger way loss (lower fitness values) will trigger an attempt to improve this by reducing speed. Apart from that, any detected violations of COLREGS (including Rule 10) will automatically trigger speed reduction, because the presence of such violations in the final set of tracks means that the method has failed to produce a safe solution.

The role of evaluation, reference tracks and reference fitness values in the evaluation process is summarised by Fig. 1, which depicts the method’s main algorithm.

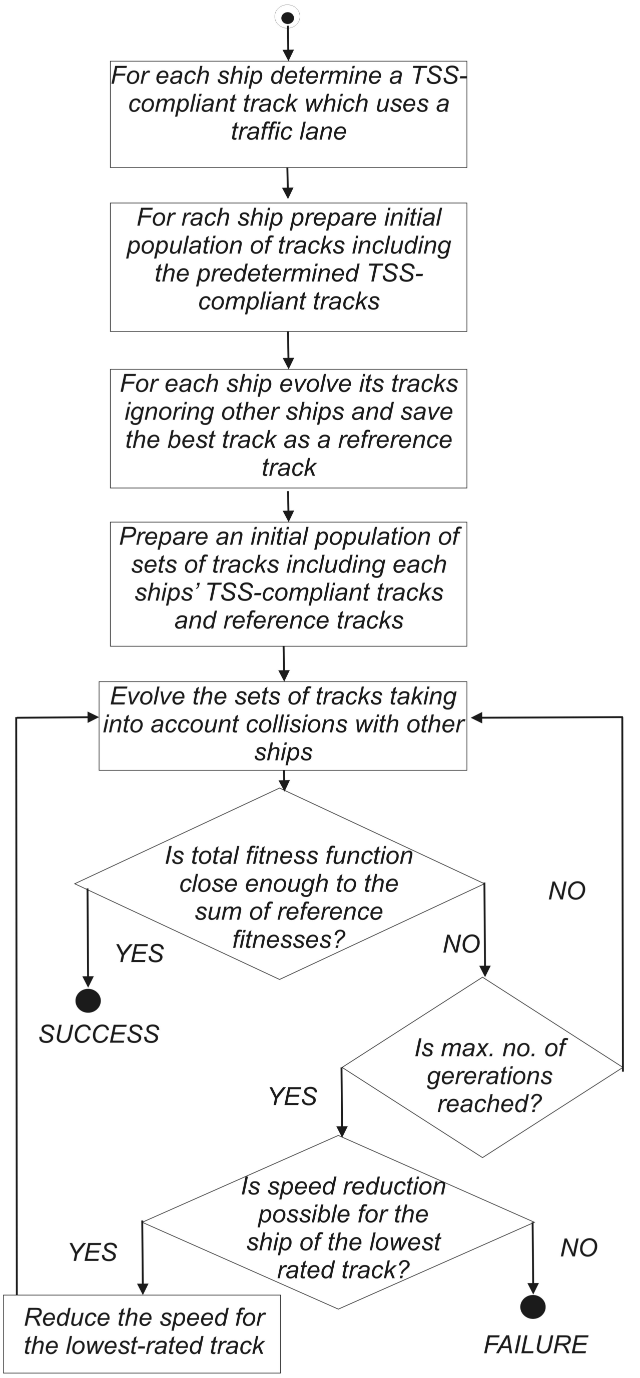

Figure 1. The ESoSST method's main algorithm including the use of reference tracks and reference fitness values.

5 FITNESS FUNCTION

The fitness function used in the method is a sum of fitness values of all trajectories in a set:

$$fitness = \sum_{i=1}^{n}\left[trajectory\_fitness_i\right], \tag{1}$$

where:

$$track\_fitness_i = \frac{basic\_track\_fitness_i}{reference\_track\_fitness_i}. \tag{2}$$

Basic track fitness is the track fitness value before normalization:

$$basic\_track\_fitness_i = track\_economy\_factor_i * \\ * scf_i * caf_i * ccf_i * tcf_i \quad . \tag{3}$$

scf_i (static constraint factor), caf_i (collision avoidance factor) and ccf_i (COLREGS-compliance factor) have already been described in the author's earlier papers. $Track\_econ\_factor_i$ is computed in a different way for ships which have to reduce their speed than for those, which do not. For changing propeller's settings, that is for ships, which reduce their speed:

$$track\_econ\_factor_i = \left(\frac{track\_length_i - way\_loss_i}{track\_length_i}\right) \tag{4}$$

For fixed propeller's settings, that is for ships which do not reduce their speed:

$$track\_econ\_factor_i = \left(\frac{track\_time_i - time\_loss_i}{track\_time_i}\right) \tag{5}$$

where:

i = the index of the current ship,

$track\_time_i$ = the total time by the i-th ship between the endpoints of its track [hours]; $time\_loss_i$ = the total time loss of the i-th ship computed as a difference between the track time and the time spent on covering a straight segment joining the track's endpoints. $time\_loss_i$ includes temporary fall in speed (for fixed propeller's settings) due to course alteration manoeuvres. $track\_length_i$ = the total length of the i-th ship's track [nautical miles]; way_loss_i = the difference between the length of the i-th ship's track and the length of a straight segment joining the endpoints of the i-th ship's track [nautical miles].

In case of speed reduction, the track economy factor would be much lower for time-oriented track economy factor (5) and would lead to unacceptably low track fitness value and total fitness function value. Therefore the length-oriented track economy factor (4) is used for ships, which reduce their speed.

tcf_i factor from formula (3) is responsible for TSS-compliance and is computed as given by (6).

$$tcf_i = \left(1 - \sum_{k=1}^{m}\left[TSS\_violation\_penalty_k\right]\right) *, \\ * \left[1 + lpf_i * \left(lef - 1\right)\right] \tag{6}$$

where m = the number of TSS rules violations registered for the current ship; k = the index of a registered violation, $TSS\_violation\_penalty_k$ = the penalty for the k-th of the registered TSS rules violations; lef = lane encouragement factor applied to encourage using traffic lanes, usually from the <1.1, 1.5> range, set to 1.2 by default; lpf_i = track's lane percentage factor (a percentage of the track's length that transits through a traffic lane).

$reference\_track\_fitness_i$ from equation (2) is the fitness value of a predetermined track of the i-th ship, found without taking into account potential collisions with other ships.

6 EXAMPLE RESULTS

In this section some examples showing the importance of a well-designed fitness function are presented. For all three scenarios it will be shown how different fitness values might be assigned to the same solution by different fitness functions and what the consequences of right or wrong evaluation are. The ship domains for all scenarios have been set to values which enable two ships only to transit through a lane parallel to each other. It has also been assumed that a ship may reduce its original speed by 0.3 or 0.5 of the initial value.

6.1 *Scenario 1*

In this scenario it will be exemplified, how fitness function leads to convergence to the right or wrong solution. The results of using fitness function without lane encouragement factor and with lane encouragement factor are shown in Figs. 2 and 3 respectively. First a fitness function with a simplified TSS-compliance factor is used, where TSS violations are penalized, but using traffic lanes is not rewarded – the formula, which is used for TSS-compliance factor, is given by (7):

$$tcf_i = \left(1 - \sum_{k=1}^{m} \left[TSS\_violation\_penalty_k\right]\right). \qquad (7)$$

As a result, a track not using a lane is chosen due to smaller way loss. This is shown in Fig. 2. In contrast to this, a full formula for TSS-compliance factor (6) is used later and the result is presented in Fig. 3. Here a track transiting through a lane is chosen, as required by COLREGS.

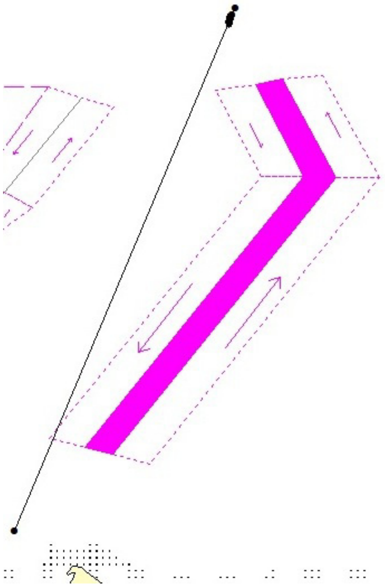

Figure 2. A ship track by a fitness function without lane encouragement factor.

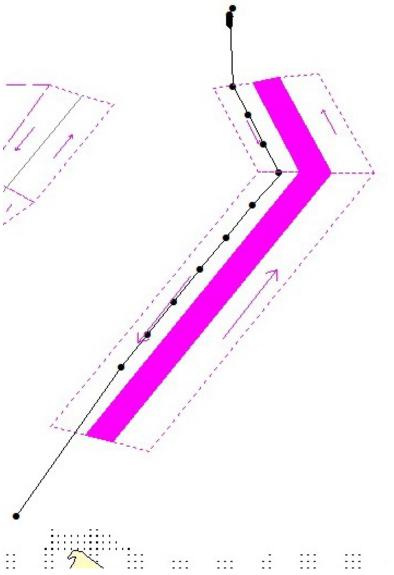

Figure 3. A ship track by a fitness function using lane encouragement factor.

6.2 *Scenario 2*

In this scenario it will be shown, how fitness function might lead to misinterpretation of the final solution. Here an incomplete (not normalized) fitness function results in accepting a solution as a near-optimal, whereas in fact it is far from optimal. The initial positions, courses and speeds of all ships are shown in Fig. 4. First, a fitness function, which does not use reference track fitness is used, with a lane encouragement factor set to 1.5. It is given by (8):

$$track\_fitness_i = track\_econ\_factor_i * \\ * scf_i * caf_i * ccf_i * tcf_i \qquad (8)$$

The solution obtained for such simplified, non-normalized fitness function is shown in Fig. 5. All three ships transit through a traffic lane (although the overtaking ship exits the lane and enters it again) and therefore, due to high lane encouragement factor, the solution is highly rated and is accepted as near-optimal, which is wrong.

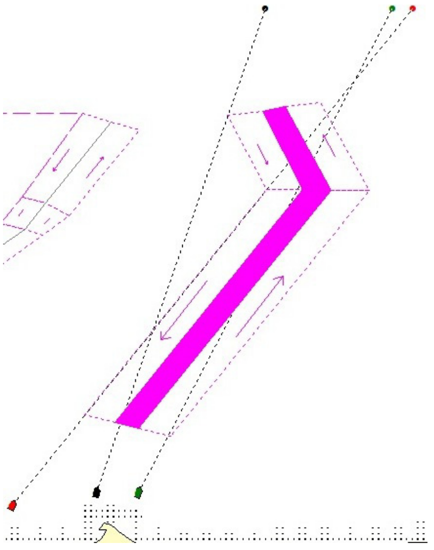

Figure 4. Initial parameters of ships from Scenario 2, speeds left to right: 15, 12 and 12 knots.

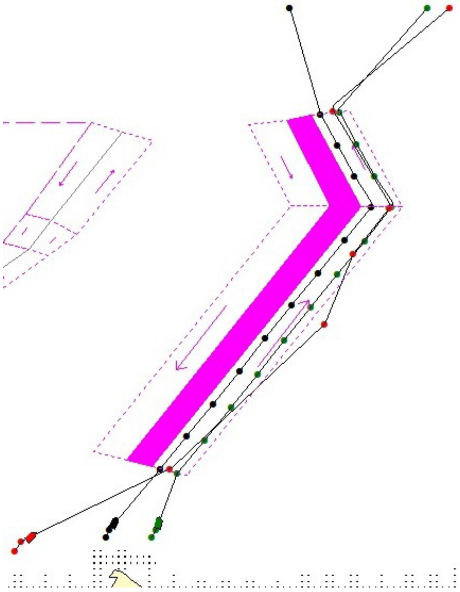

Figure 5. A solution accepted by the method with a simplified, non-normalized fitness function

In contrast to this, when evaluated according to (2) and (3) the fitness value of the solution from Fig. 5 is much lower, because the reference fitness value for this track is high (the reference track is shown in Fig. 6). Therefore the solution from Fig. 5 is dismissed and the method applies a speed reduction for this ship. The ship reduces its speed by 0.3 of the original value (from 15 to 10.5 knots). The final solution is shown in Fig. 7. The two ships having the same speed transit through a lane parallel to each other, while the third ship reduces its speed and follows them keeping a safe distance.

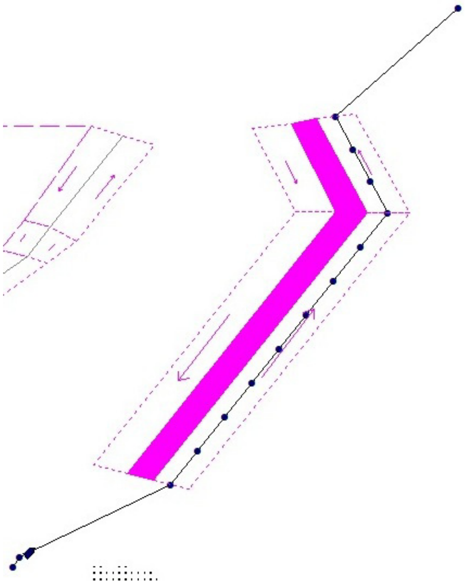

Figure 6. A reference track of the fastest ship from Scenario 2 (far left in Figure 5)

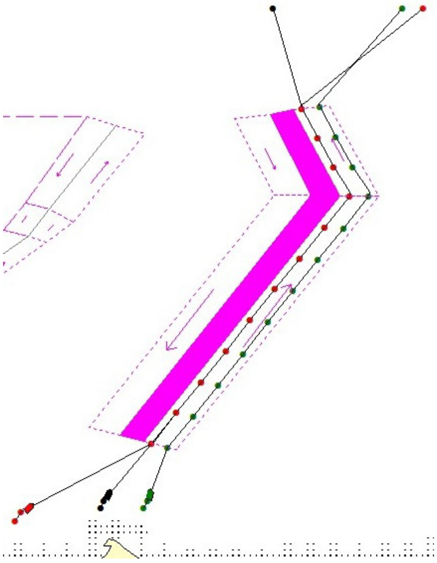

Figure 7. A solution returned by the method with normalized fitness function – the fastest ship reduces its speed

6.3 *Scenario 3*

In this scenario it will be shown, how incomplete fitness function results in another kind of misinterpretation of the final solution. Here a simplified, non-normalized fitness function leads to a situation opposite to that from Scenario 2 – to dismissing an acceptable solution and to undesired and unsuccessful attempts to improve it by applying speed reduction manoeuvres. The initial positions, courses and speeds of all ships are shown in Fig. 8. First, a simplified fitness function, which does not use reference track fitness is used, similarly to Scenario 2. As a result, the solution shown in Fig. 9 is underrated and dismissed as unacceptable.

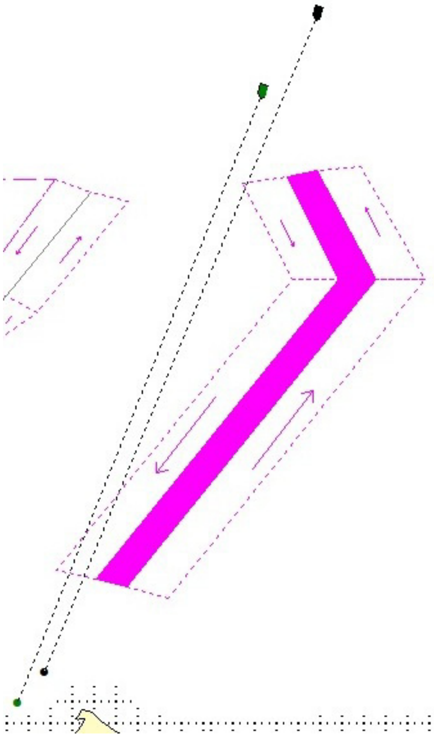

Figure 8. Initial parameters of ships from Scenario 5, speeds left to right: 15 and 18 knots

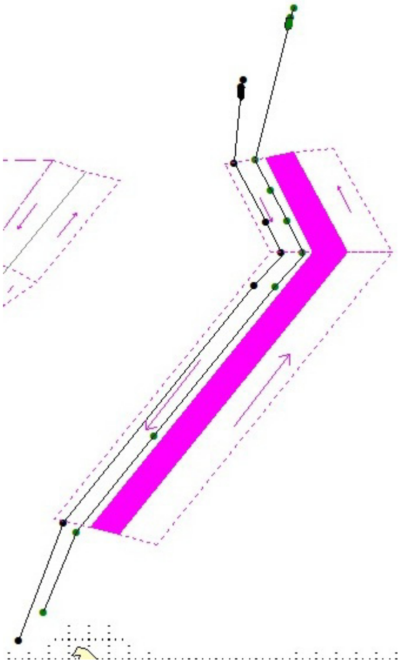

Figure 9. A solution dismissed by the method with a simplified, non-normalized fitness function due to seemingly too large way loss

Because of the lack of COLREGS violation, the method assumes that the solution's low fitness value is a consequence of way loss due to collision avoidance and tries to improve the solution by applying speed reduction manoeuvres. These manoeuvres however obviously cannot decrease way loss and the method eventually returns a similar solution (shown in Fig. 10) accompanied by a message that no acceptable solution could be found for the given parameters. In contrast to this, when evaluated according to formulas (2) and (3) the fitness value of the solution from Fig. 9 is properly rated, based on the reference track (shown in Fig. 11). Therefore the solution from Figure 9 is accepted as a correct one, as it should.

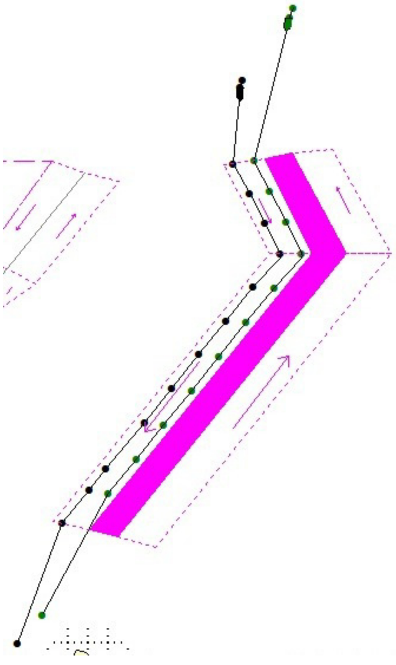

Figure 10. A solution dismissed by the method with a simplified, non-normalized fitness function after a failed attempt to decrease way loss by speed reduction

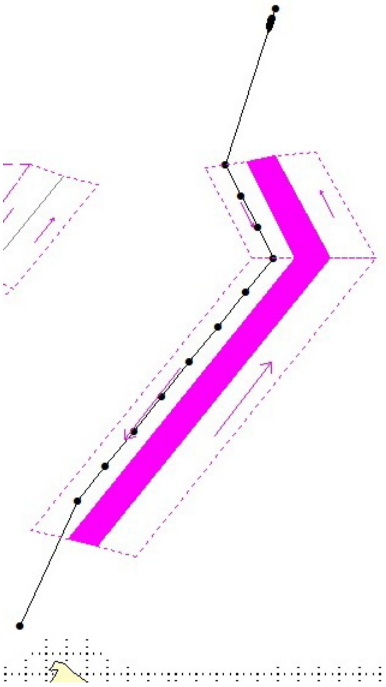

Figure 11. A reference track of the faster of the two ships in Scenario 3.

7 SUMMARY AND CONCLUSIONS

In the paper it has been discussed how the evaluation phase affects the results returned by the Evolutionary Sets of Safe Ship Trajectories method. In particular, it was shown that the fitness function is crucial not only for the convergence to the acceptable solution but also for the assessment of this solution's quality and telling when it actually is acceptable. The imprecise assessment of the already found ship track can lead to communicating a failure or to planning a speed reduction maneuver when it is neither necessary nor desired. As has been shown by the examples, the fitness function presented in the paper successfully deals with the above mentioned issues and results in planning ship tracks which are safe, economical and TSS-compliant.

As has already been mentioned in Section 2, the current version of the method assumes good visibility and does not support Rule 19 of COLREGS. Also, as for now, the method does not take into account the bathymetry data and ship's draught. Additionally, due to the fact, that no information on ships maneuverability is provided by AIS, the method applies a largely simplified ship dynamics model. Therefore further research on the ESoSST method and further development of the method is planned. It will be focused on handling all the above mentioned problems. Future tasks also include working on the scalability of the presented approach – dealing with larger numbers of ships and larger areas.

REFERENCES

Anwar N. & Khalique A. 2006. *Passage Planning Principles.* Witherby Seamanship International, Livingston, United Kingdom

Cockcroft A.N. & Lameijer, J.N.F. 2011. *A Guide to Collision Avoidance Rules*. Butterworth-Heinemann.

Coldwell T.G. 1983. Marine Traffic Behaviour in Restricted Waters. *The Journal of Navigation*. 36, 431-444.

COLREGS. 1972. [with amendments adopted from December 2009]). *Convention on the International Regulations for Preventing Collisions at Sea*. International Maritime Organization, London.

Lisowski J. 2007. The Dynamic Game Models of Safe Navigation, *International Journal on Marine Navigation and Safety of Sea Transportation*, Vol. 1, no. 1., 11-18

Michalewicz Z. & Fogel, D.B. 2004. *How To Solve It: Modern Heuristics.* Springer-Verlag.

Smierzchalski R. & Michalewicz, Z. 2000. *Modeling of a Ship Trajectory in Collision Situations at Sea by Evolutionary Algorithm, IEEE Transactions on Evolutionary Computation* No. 3 Vol. 4, pp. 227-241.

Szlapczynski R. 2011. Evolutionary Sets of Safe Ship Trajectories: A New Approach to Collision Avoidance. *The Journal of Navigation*. 64, 169-181.

Szlapczynski R & Szlapczynska J. 2011, *On Evolutionary Computing in Multi-Ship Trajectory Planning. Applied Intelligence*, vol. 37, iss. 2, 155-174, Springer Netherlands, http://www.springerlink.com/content/j40qp07031634568 (free access)

Szlapczynski R. 2012. Evolutionary Sets of Safe Ship Trajectories within Traffic Separation Schemes. *The Journal of Navigation*. vol. 66, iss. 1., pp. 65-81.

Szlapczynski R & Szlapczynska J. 2012. Evolutionary Sets of Safe Ship Trajectories: Evaluation of Individuals. *International Journal of Marine Navigation and Safety of Sea Transportation* Vol. 6, No. 3, 2012, s. 345-353.