A method for ice-aware maritime route optimization

Robert E. Guinness, Jarno Saarimäki, Laura Ruotsalainen, Heidi Kuusniemi
Department of Navigation and Positioning
Finnish Geodetic Institute
Kirkkonummi, Finland
firstname.lastname@fgi.fi

Floris Goerlandt, Jakub Montewka
Department of Applied Mechanics
Aalto University
Espoo, Finland
firstname.lastname@aalto.fi

Robin Berglund, Ville Kotovirta
Department of Solutions for natural resources and environment
VTT Technical Research Centre of Finland
Espoo, Finland
firstname.lastname@vtt.fi

***Abstract*—We present a method for *ice-aware* maritime route optimization. Our aim is to increase the safety and efficiency of maritime transport under icy conditions. The proposed method is based on the A* algorithm, developed by Hart et al. It uses a model of maritime navigation, consisting of (1) a sea spatial model, (2) ship maneuverability model, (3) sea ice model, and (4) ship performance model. The sea ice model, which provides a snapshot of the sea ice conditions, is based on previous work by the Finnish Meteorological Institute. The ship performance model, based on previous work by Kotovirta et al., estimates ship transit speed as a function of ice conditions and ship design parameters. The main novelties in this paper are the application of the A* algorithm to maritime route optimization and development of an associated cost function that takes into account ice conditions and available icebreaker assistance. We present preliminary results based on the method, using the Baltic Sea as a case study. Generated routes are compared with historical routes under the same ice conditions. Areas of future work and needed enhancements are briefly discussed.**

Keywords*— routing; ice navigation; maritime navigation; route optimization

I. Introduction

Conventionally, navigation and routing functions of vessels are the responsibility of the vessel's captain and his or her designated crew, who rely on their extensive experience and training to navigate the vessel safely and efficiently. Increasingly, various data sources, such as electronic charts, radar imagery, and Automatic Identification System (AIS) information, are used to assist the captain in planning the vessel's route and in real-time navigation to adjust the route based on the evolving conditions at sea. In challenging maritime environments, such as in winter conditions at high latitudes, these various data sources play a critical role [1]. As the demand for maritime transport increases, especially in northern and Arctic regions, the need for systems that specifically address the problem of ice navigation is quickly rising.

Despite the great success of automated route planning and navigation for land vehicles, such as car navigators or online driving directions services, maritime navigation and route planning remains a mostly manual task, which requires extensive training and great vigilance to be carried out safely and efficiently. Various software systems have been developed to assist the crew in routing and navigation functions, such as the Jeppesen Vessel and Voyage Optimization Solution (VVOS), as well as onboard Decision Support Systems (DSS), which mostly aim to assist the captain in mitigating risks to the vessel [2]. A large gap still remains, however, between the available information sources and the ability to utilize them in an efficient manner. Previous research has attempted to narrow this gap, most notably the European project, Ice Ridging Information for Decision Making in Shipping Operations (IRIS) [3]. Most of the current commercial solutions and systems, however, are not applicable or optimized for navigation in ice-covered waters.

Furthermore, there are many documented instances where poor navigation choices at sea, either due to lack of available information or poor training, have led to suboptimal routes or even unsafe routes to be used in ice-covered waters. For example, in March 2010 around fifty vessels were stuck in Baltic Sea ice, including six passenger ships, requiring extensive icebreaker assistance to free the stuck ships. During this episode a large passenger ship, MS Amorella, collided with another passenger ship, MS Finnfellow [4]. Countries with high latitude sea borders, such as Finland, Sweden, and Russia, expend significant resources to maintain and operate icebreaker fleets, in order to assist vessels that become stuck in sea ice, as well as to produce channels in the ice for easier transit (see e.g. [5]).

The goal of the present work is develop a method that takes into account available information about sea ice conditions and icebreaker operations, in order to determine optimal routes, both in terms of minimizing crossing time and minimizing risk of ships getting stuck or damaged due to severe ice conditions. Due to the Baltic Sea being one of the busiest waterways in the world that experiences significant ice coverage every winter, this paper will focus on examples from this region.

The paper is organized as follows: We first review related work especially in the areas of maritime route optimization and ice navigation. In Section III, we describe (1) the model we used for the sea environment, including a model of sea ice conditions, (2) the assumptions we made about how ships can maneuver in the sea environment, and (3) our model of ship transit speeds as a function of ice conditions. In Section IV, we describe icebreaker operations, in order to elucidate the role they play in route planning. Section V presents the route optimization method, which utilizes all the elements discussed

This work was supported in part by the Finnish Funding Agency for Innovation (Tekes) under the ARCSAT project.

978-1-4799-3320-4/14/$31.00 © 2014 IEEE

in Sections III-IV. In Section VI, we discuss results of the route optimization method, including preliminary validation of the results based on historical route information. Finally, we conclude our work in Section VII, summarizing the main benefits of our proposed method and describing areas of future work that still remain.

II. Related Work

Research related to maritime route optimization can be traced back primarily to a seminal 1960 paper by Hanssen and James, which describes a system developed by the United States Hydrographic Office to optimize transoceanic crossings, utilizing predictions of wind, waves and current [6]. In the years following, many different route optimization methods have been proposed, however, few of them are applicable to navigation in ice-covered waters. One notable exception is [3]. This method described in this paper employed a cost function based on ship speed and navigation restrictions, such as fairways and shallow water. The authors' ship speed estimation technique took into account ice conditions based on three ice parameters, including level ice thickness, ridged ice thickness, and ice concentration. They evaluated the use of three different optimization techniques for minimizing this cost function, in order to compute an optimal route, including Powell's method, the polytope method, and simulated annealing. The method the authors chose for prototype implementation (Powell's method), however, cannot guarantee to find a global minimum of the cost function [7]. In other words, an initial guess of the route is supplied, and the method will find a local minimum of the cost function with respect to the initial guess.

III. Development of a Maritime Model

The purpose of this section is to describe the model of the sea and ship motion that was used in this study. The model can be divided up into four parts: (1) sea spatial model, (2) ship maneuverability model, (3) sea ice model, and (4) ship performance model.

A. Sea Spatial Model

The main difference between a car navigation application and a maritime navigation application is that a car is assumed to only travel on roadways, whereas a ship can generally travel anywhere where the sea depth is greater than the ship's draught and in any direction within this area, provided only that it stays within such area. This degree of freedom makes the route planning problem more complex because there is an infinite number of possible routes between any two given points. Despite this fact, we note that similar routes will have similar crossing times. Thus, we can divide the sea area into a finite number of points and use this set of points as an approximate model for the actual sea.

Let S represent a set of points in a discrete model of the sea and $m = |S|$ represent the number of points in this set. For each point $p \in S$, there will be associated information, such as its latitude and longitude, the sea depth at that point, as well as ice-related information, as will be discussed below. The complete set of points S and this associated metadata are the primary components of the model of the sea. We refer to this part of the modeling process as a *discretization of the sea*.

There are a number of different ways to discretize a given sea area, but in this work we chose the most simple approach, which is to define a uniform grid of points with a suitable grid spacing covering the whole sea area. The drawback of this approach is that if the grid spacing is small (i.e. a dense grid of points), there will be a large number of points, which leads to more computation required for the routing algorithm. If the grid spacing is large, then the required computation will be less, but the resulting model will not accurately represent the actual sea. That is, there will only be a limited number of positions and directions in which the ship can travel. Fig. 1 shows an example grid applied across the Baltic Sea, where the grid points are shown in red. In this example, the grid spacing is relatively sparse (~20 km between each grid point). The actual grid spacing used in this work was approximately 1 km. This value, however, can be adjusted according to the requirements of the end user.

Also shown in Fig. 1 is the concept of a *depth mask*, which is a binary image representing where a ship can and cannot travel based on the depth of the sea. If a given ship has a draught of r, then the value of the depth mask is defined at each coordinate in the image as follows:

$$D_{ij} = \begin{cases} 1 & \text{for } d_{ij} > r \\ 0 & \text{for } d_{ij} \le r \end{cases} \tag{1}$$

where D_{ij} is the value of the depth mask at coordinate (i, j), d_{ij} is the depth of sea at coordinate (i, j), and r is the draught of the ship. If desired a safety factor can be applied, such that:

$$r = (1 + \varepsilon) r_{\text{actual}} \tag{2}$$

where r_{actual} is the ship's actual draught and ε is a positive value representing the safety factor. In Fig. 1, a value of r = 10m was used. The white regions represent the areas where the depth mask has a value of 1, and the black regions represent a value of 0.

Note that because the depth mask is a function of a ship's draught, the resulting model of the sea is customized to a specific ship or class of ships. In practice, any implementation of this method should pre-compute a set of depth masks appropriate for the set of ships for which the routing software will be used. Thus, the software could apply the most suitable depth mask as a lookup table prior to start of route optimization.

B. Modeling Ship Maneuverability

To model how a ship can move among the grid points described above, we define a set of neighbors $\mathcal{N}$ and assume a ship can move from its current position (i.e. grid point) to one of its neighboring grid points defined by this set. For example, if the set of neighbors were defined as $\mathcal{N}$={north, south, east, west}, then the set of vectors from the current position to each of the neighbors (which we refer to as the set of direction vectors $\mathcal{V}$) would form the four cardinal compass directions. As a result, the ship could only move in these four cardinal directions. Since this is obviously not sufficient to accurately describe the real maneuverability of a ship, additional neighbors must be added to the set.

Fig. 1. Example of a depth mask for the Baltic Sea with a threshold value *r* set to 10 m. The black regions represent unnavigable areas. The red dots represent nodes in the discrete model of the sea. In the figure, a node spacing of about 20 km is shown, but in practical implementations a smaller spacing is generally used.

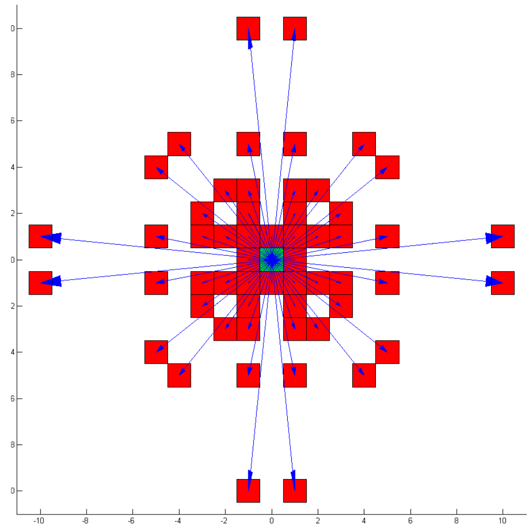

Fig. 2. The set of neighbors $\mathcal{N}$ and direction vectors $\mathcal{V}$ used to model ship maneuverability.

Let $n = |\mathcal{N}| = |\mathcal{V}|$ represent the number of neighbors defined in the ship maneuverability model. In this work, we used $n = 56$ and defined the sets $\mathcal{N}$ and $\mathcal{V}$ as shown in the Appendix (i.e. Table III)[1]. Note that because a uniform rectangular grid of points was used, the length of each of the vectors in set $\mathcal{V}$ is not necessarily the same. In fact, in defining these sets, the aim was to keep the difference in compass directions between the direction vectors as close to equal as possible. Fig. 2 illustrates the sets $\mathcal{N}$ and $\mathcal{V}$ defined and used in this work, where the red squares represent the neighbors (relative to the central square drawn in green) and the blue arrows represent the direction vectors. The average angle between adjacent direction vectors is 6.43° with a minimum separation of 4.97° and maximum of 8.13°.

Thus, the set defined by $\mathcal{N}$ represents a discretization of the maneuverability of a ship at sea, similar to how the set of grid points S discretized the sea into a finite set of points. The choice of n is a tradeoff between having an accurate model and the computational complexity of the routing algorithm which will use the model. Similar to the choice of grid spacing, n can be adjusted according to the requirements of the end user.

Together the sets S and $\mathcal{N}$ constitute a graph, consisting of m nodes and a set of edges between the nodes, defined according to $\mathcal{N}$. In other words, $\mathcal{N}$ determines the connectivity of the graph. Because nodes on the edge of the sea boundary do not necessarily have n valid neighbors, the exact number of edges depends on the dimensions of the sea area. The upper bound for the number of edges, however, is m*n. Fig 3. shows a portion of the graph structure used in this work (i.e. n = 56). The blue lines show the edges drawn between nodes in the graph (which are shown in red). The figure has been zoomed in to cover only nine points because when zoomed further out, it quickly becomes difficult to discern the edges separately.

[1]Note that the set $\mathcal{V}$ is uniquely defined by the set $\mathcal{N}$. As such, a reference to $\mathcal{V}$ implies the corresponding set $\mathcal{N}$ and vice versa.

C. Modeling Sea Ice

The ice data were obtained from the hindcasts generated using the HELMI multicategory sea-ice model, developed at the Finnish Meteorological Institute; for details see [8]. The model resolves ice velocity, internal ice stress, ice concentration and ice thickness. Thickness is resolved for seven categories: five level ice categories, plus rafted ice and ridged ice. The ice model is discretized in a curvilinear coordinate system called a c-crid, a common solution when there are both fields of velocities and velocity-dependent properties to be solved. The grid has 415 nodes from west to east and 556 nodes from south to north. The SW lower corner coordinates are 56.74° N 16.72° E, NE corner coordinates 65.99° N 30.48° E, and the increment is 1/30 degrees eastwards and 1/60 degrees northwards. This is approximately 1 NM in both directions at 60°N.

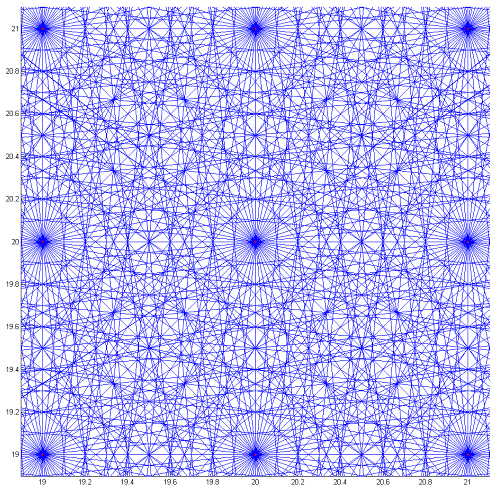

Fig. 3. The blue lines represent connections between nodes which are valid paths for a ship to travel. Only nine nodes are shown in the figure, whereas in an area the size of the Baltic Sea, there may be tens of thousands of nodes, depending on the chosen grid size.

Ice motion is determined by the momentum balance equation, which takes into account the Coriolis force, wind and water stresses, sea surface tilt term and internal friction of ice, which is the divergence of internal stress tensor. The magnitude of internal friction is used as the principal model variable to describe compression. It is to be noted that the viscous-plastic rheology does not describe elastic stresses and the internal stress arises from the interactions of moving ice. Forces arising in a static ice are included by assuming a negligibly slow viscous creep. Roughly, the internal friction term can be interpreted to describe the forces arising when ice floes are pushed and sheared against each other, or broken and heaped into ridges. Thus it is a good descriptor for the interaction between dynamical ice cover and an ice-going ship. This is manifested as ice forces against the ship hull and as the closing of channels, or other phenomena that navigators associate to compressive ice conditions. The internal friction magnitude has typical values ranging from 0 to 10 N/m^2. The magnitude acts as a proxy for ice compression, scaled to semi-empirical compression numeral 0-4, where 0 means no compression and 4 stands for extreme severe compression, see Table I. However, to estimate the actual local forces additional scaling arguments must be taken into account such as floe size and other ice cover geometry.

TABLE I. ICE COMPRESSION CONVERSION

Ice pressure obtained from the HELMI model [10^4 N]	Practical scale	
0 – 500	no significant compression	0
500 – 1000	mild pressure	1
1000 – 2000	moderate pressure	2
2000 – 3000	severe pressure	3
>3000	extreme severe pressure	4

D. Modeling Ship Performance in Ice

This chapter provides an outline of the adopted approach for estimating ship performance in ice conditions, which is used in the cost function of the path planning algorithm.

Various advanced theoretical approaches have been developed to estimate vessel performance in ice conditions, including analytical models [10-11] and simulation-based calculation procedures [12-13]. These typically require detailed information regarding the ship hull design, while being limited in application in real-world ice conditions as the mentioned models only consider level ice. Methods based on full-scale onboard data have been developed [14] and recently, a probabilistic model based on a combination of data from the Automatic Identification System (AIS) and ice information has been developed [15, 23].

In the current approach to vessel performance, an approach presented in Kotovirta et al. is applied [3]. While it has the limitation of only accounting for level ice, ice ridges and pack ice, its computational efficiency is beneficial in the context of route optimization. The methodology is based on a semi-empirical model by Lindqvist [16], as modified in La Prairie et al. [17] and Riska et al. [18]. In ridged ice fields, the ship transit speed $v_{i,eq}$ (m/s) is determined by numerically solving following force balance equation, where ice resistance and propulsion power are in equilibrium:

$$T_{\mathrm{net}}(v) = R_{\mathrm{tot}}\left(h_{\mathrm{i}}, h_{\mathrm{eq}}, v_{i,\mathrm{eq}}\right) \tag{3}$$

Here, T_{net} is the net available trust (kN), R_{tot} the total ice resistance (kN), h_i the level ice thickness (m) and h_{eq} the mean thickness of the ice rubble (m). h_i and h_{eq} are derived from the sea ice model (see Section IIC). The net available trust T_{net} is obtained from following simplified formula, see [19]:

$$T_{\mathrm{net}}(v) = \left(1 - \frac{1}{3}\frac{v}{v_{\mathrm{ow}}} - \frac{2}{3}\left(\frac{v}{v_{\mathrm{ow}}}\right)^2\right) T_{\mathrm{pull}} \tag{4}$$

Here, v is the average ship speed (m/s), v_{ow} the ship speed in open water (m/s) and T_{pull} the bollard pull force (kN), i.e. the available trust when a vessel is stationary with engines running:

$$T_{\mathrm{pull}} = K_{\mathrm{e}}\left(P_{\mathrm{s}} D_{\mathrm{p}}\right)^{2/3} \tag{5}$$

with propulsion power P_s, propeller diameter D_p and quality efficient for bollard pull K_e.

While not fully physically correct, the total ice resistance R_{tot} is taken as a sum of resistance components stemming from breaking through level ice R_i and passing through ridge rubble R_r:

$$R_{\mathrm{tot}}(v) = R_{\mathrm{i}}(h_{\mathrm{i}}, v) + R_r\left(h_{\mathrm{eq}}, v\right) \tag{6}$$

Riska et al. [18] assume a linear relation for the relation between level ice thickness and speed:

$$R_{\mathrm{i}} = C_1 + C_2 v \tag{7}$$

with:

$$C_1 = f_1 \frac{1}{2\frac{T}{B}+1} B L_{\mathrm{par}} h_{\mathrm{i}} + (1 + 0.021\varphi)\left(f_2 B h_{\mathrm{i}}^2 + f_3 L_{\mathrm{bow}} h_{\mathrm{i}}^2 + f_4 B L_{\mathrm{bow}} h_{\mathrm{i}}\right) \tag{8}$$

and:

$$C_2 = (1 + 0.063\varphi)\left(g_1 h_{\mathrm{i}}^{1.5} + g_2 B h_{\mathrm{i}}\right) + g_3 h_{\mathrm{i}}\left(1 + 1.2\frac{T}{B}\right)\frac{B^2}{\sqrt{L}} \tag{9}$$

The resistance in ridges R_r is calculated as:

$$R_{\mathrm{r}} = 4 C_3 h_{\mathrm{eq}}^2\left[B + 2 C_{\psi} h_{\mathrm{eq}}\right]\left(\mu_{\mathrm{h}} \cos\varphi_2 + \sin\psi \sin\alpha_2\right) + C_4 L_{\mathrm{par}} h_{\mathrm{eq}}^2 + C_5\left[\frac{LT}{B^2}\right]^3 h_{\mathrm{eq}} A_{\mathrm{wf}} F_{\mathrm{n}}^2 \tag{10}$$

Here, μ_h is the friction between hull and ice, φ_2 the angle with buttock line at B/4 with the horizontal, α_2 the waterline entrance angle at B/4, L_{par} the length of the parallel midbody, L, B and T the vessel length, width and draft, A_{wf} the bow waterplane area, and F_n the Froude number. All necessary constants are given for the considered vessel in Table II.

The effect of ice concentration on ship performance in ice, focusing on attainable speed and hull loads, has been studied using a simulation model of ship dynamics in ice, applying analytical formulae for ice resistance [20]. The adopted approach in the route optimization is based on empirical formulae and a set of simplifying assumptions, see [3].

TABLE II. SHIP PERFORMANCE CONSTANTS

Constant	Value	Unit
f_1	230	N/m^3
f_2	4580	N/m^3
f_3	1470	N/m^3
f_4	290	N/m^3
g_1	18900	$N/(m^{2.5}/s)$
g_2	670	$N/(m^3/s)$
g_3	1550	$N/(m^{3.5}/s)$
C_3	850	N/m^3
C_4	42	N/m^3
C_5	1300	N/m^3
C_ψ	$0.047\psi - 2.115$	-
L	193.7	m
L_{par}	136	m
B	30.2	m
T	12.0	m
μ_h	0.05	-
φ_2	0.7156	rad
α_2	0.6981	rad
A_{wf}	496.79	m^2

With an ice concentration C less that C_0, the vessel is assumed to manoeuver around ice floes and sail at the open water speed v_{ow}. For ice concentrations higher than C_1, the ship operates at $v_{i,eq}$, the speed through the ridged ice field. Between C_0 and C_1, the resulting transit speed v_{tr} is assumed as a linear combination of v_{ow} and $v_{i,eq}$:, assuming C_0=70% and C_1=95%:

$$v_{\mathrm{tr}} = \begin{cases} v_{\mathrm{ow}} & C \le C_0 \\ \frac{(C_1 - C)v_{\mathrm{ow}} + (C - C_0)v_{\mathrm{i,eq}}}{(C_1 - C_0)} & C_0 < C < C_1 \\ v_{\mathrm{i,eq}} & C \ge C_1 \end{cases} \qquad (11)$$

While the vessel performance is largely based on semi-empirical formulations and does not account for effects of compression, which can be significant [21], the approach leads to a reasonable estimate of total transit time [3, 17].

IV. WINTER MARITIME SHIPPING IN THE BALTIC SEA

In spite of the sea ice in wintertime, most of the ports around the Baltic Sea are kept open all year around for shipping. This is possible thanks to specially built icebreakers (8 Finnish, 9 Swedish, 2 Estonian, 3 larger and several smaller Russian icebreakers operate in the Northern Baltic Sea) that assist the merchant ships through the ice field.

Cooperation between the nations around the Baltic Sea enables more efficient use of the icebreakers, as many ship routes have common parts irrespective of the port of destination. Especially in the Bay of Bothnia and the Sea of Bothnia, the Finnish and Swedish icebreakers organize their assistance activities as a joint fleet and the operating costs are shared between the Finnish and Swedish governments based on actual assistance provided and the port of destination of the assisted ships.

To ensure efficient and safe navigation, the ships bound for ports surrounded by ice have to fulfill some minimum requirements regarding capability to navigate through the ice, both from a safety (e.g. hull strength) and efficiency point-of-view (e.g. machine power). These requirements have been formulated into ice class rules defined by classification societies. The Ice Class notation given by a Classification society can be mapped to an Equivalent Finnish/Swedish ice class using mapping tables issued by the Finnish and Swedish authorities. The port-specific ice class restrictions (Ice Restrictions for short) are determined by the maritime Administration in the country of the destination port and are updated depending on the ice conditions and forecasts. These ice restrictions are published on the Internet and also included in the daily ice charts issued by the national Ice Services.

A ship coming from the southern part of the Baltic Sea, bound for a Swedish or Finnish port having a valid ice restriction, is obliged to contact a coordinating point announcing, among other things, the port of destination of the ship. The coordinating point then checks that the ship fulfils the ice class requirements and tells the ship the coordinates of the next ice waypoint and also which icebreaker to contact. Often, when the ship is considered capable of navigating independently for part of the voyage (not necessarily having to be escorted by an icebreaker), the ship is given several ice waypoints defining a coarse route through the ice field. The details on how to navigate through the ice field between the ice waypoints, which may be separated by tens of nautical miles, are left to the ship itself. The waypoints are often communicated to the ship over VHF, but sending the ice waypoints in digital form to the ship using e-mail or an AIS addressed message, is becoming more and more common. Icebreakers give lowest priority to ships which do not follow these waypoints, the result of which is that ships in general follow them whenever there is a significant chance of requiring icebreaker assistance.

The ice waypoints are determined by the icebreakers and registered in the Icebreaker information system, IBNet, maintaining the overall set of valid points. The waypoints are updated when needed as a consequence of changing ice conditions. Some of the waypoints may be valid for weeks, but often they are adjusted with 2 -3 days intervals. Satellite images (especially Synthetic Aperture Radar images), ice drift forecasts, ship speed monitoring (i.e. how ships have been able to proceed through the ice field) and ice observations from other icebreakers are used as background information. This information is then combined by the experienced icebreaker captain into a set of waypoints that are registered in the IBNet system and communicated to the ships. A tool that could simulate different routing alternatives for ships with varying capabilities could help the icebreakers to determine and adjust the ice waypoints as part of an optimal (safe and efficient) winter navigation system.

V. Route Optimization

The classic algorithm for route optimization, also known as pathfinding, using graph structures is called A* [22]. A* finds an optimal path according to a *cost model* and subject to the constraints imposed by the graph structure. In our case the graph structure includes the set of points S resulting from the discretization of the sea, as described in Section IIIA, plus the set of edges connecting the grid points, as described in Section IIIB.

For the details of the A* algorithm, we refer to reader to [Hart et al., 1968]. In short the algorithm begins at the origin point, computes the cost to travel to that point's neighbor nodes (according to the cost model), and then sorts these neighbor nodes in priority order according to the following heuristic:

$$F_{ij} = G_{ij} + H_{ij} \tag{12}$$

where F_{ij} is the value used for sorting, G_{ij} is the cost to get to the node with coordinates (i, j), and H_{ij} is a heuristic estimate of the cost to from node (i, j) to the destination. After computing the value of F_{ij} for each neighbor node, the algorithm chooses the lowest F_{ij} among all the nodes (i, j) on the so-called "open list" and then computes F_{ij} for the neighbor nodes with respect to the chosen point (i, j), known as the current parent node.

The algorithm repeats this process recursively until the destination node has been reached. As each neighbor node is "visited" by the algorithm, the cost to get to that node from the current parent node must be saved (i.e. G_{ij}), as well as the coordinates of that parent node. Because most nodes have multiple parent nodes from which they can be visited, occasionally the algorithm will revisit nodes already on the open list. As G_{ij} is recalculated, if it is less than the stored value of G_{ij}, the old value as well as the coordinates of the parent node are replaced with the current values. This is because this new path to the node (i, j) represents the shortest path to that node among all attempted paths. In this way, the algorithm is storing a set of candidate sub-paths between the origin and the visited nodes in S. After the algorithm reaches the destination node, it can trace back through the stored parent nodes to find the best overall path between the origin and the destination.

Next, we define the cost model used by the function G_{ij}. It is a function of both the geometry (i.e. distance between nodes in the graph), the ice conditions (i.e. estimated speed that the ship can travel under given ice conditions, as derived in Section IIIC), and the distance to the nearest icebreaker waypoint (as discussed in Section V), namely:

$$G_{ij} = \sum_{k=\text{origin}}^{(i,j)} \left(\frac{d_{k\to\text{next}}}{v(k, d_{\text{IB}})} \right) \tag{13}$$

representing a Riemann sum estimating the time it takes to travel from the origin to the destination; (i,j) are the coordinates of the point immediately before the destination point; the summand contains $v(k, d_{ib})$, the estimated ship speed at point (a, b) along the route and at distance d_{IB} from the icebreaker waypoints, and $d_{k\to\text{next}}$, the length of the line segment between point k to point k+1. The function $v(k, d_{\text{IB}})$ is defined as follows:

$$v(k, d_{\text{IB}}) = \begin{cases} v_{\text{tr}}(k) & \text{for } d_{\text{IB}} > c \\ v_{\max} & \text{for } d_{\text{IB}} \le c \end{cases} \tag{14}$$

where v_{tr} is defined according to Eq. 11 and $v_{\max}$ is the maximum transit speed, normally corresponding to nominal speed of the modeled ship in open water conditions. To determine this value at any given point (a, b), we linearly interpolated the two-dimensional grid of ship transit speeds described in Section IIIB. d_{ib} is defined as the shortest distance between the point (a, b) and the line formed by the icebreaker waypoints, and c is a threshold value set to control how close to the icebreaker waypoints the ship should come before it is to receive icebreaker assistance. In this study, we used a value for c of approximately 2 km.

The remaining element to define is the heuristic function H_{ij}. We used the following:

$$H_{ij} = \frac{\sqrt{(x_i - x_{\text{dest}})^2 + (y_j - y_{\text{dest}})^2}}{v_{\max}} \tag{15}$$

where x_i is the x-coordinate of the node (i, j), x_{dest} is the x-coordinate of the destination, y_j is the y-coordinate of the node (i, j), y_{dest} is the y-coordinate of the destination, and $v_{\max}$ is the maximum. We used a local Cartesian coordinate system in order to speed up computation.

VI. Validation of Results and Discussion

Although a ship's crew has great freedom in choosing its route to their intended destination, in practice, especially in the relatively narrow Baltic Sea, ships follow more restricted pathways. Under ice conditions, ships need to follow icebreaker instructions (i.e. waypoints), and thus they have even less route options. Despite these constraints, there is much to be gained by optimizing the route based on the best available knowledge about the ice conditions. We see that the Baltic Sea area is ideal for development and testing of *ice-aware* route optimization, as it is one of the busiest ice covered sea areas, and therefore there are statistical data about ship transits available for validation, as well as model and observational data about the ice conditions.

Nonetheless, one of the challenges in this research topic is the difficulty of validating whether the chosen route optimization algorithm actually produces optimal results. That is, if the goal is to find a route that minimizes the crossing time, how can one show that the generated route gives the minimum time among all possible routes? Earlier work has shown that if a well-defined (i.e. admissible) heuristic H_{ij} is used, then A* will find the globally optimal path between the origin and destination [24]. H_{ij} is considered admissible as long as it never overestimates the actual cost to get to the destination from a given node (i, j). Because the Euclidean distance is always shorter than the geodetic distance and because the ship can never travel faster than $v_{\max}$, H_{ij} is a conservative estimate of the shortest crossing time between the node (i, j) and the destination. For this reason, H_{ij} always underestimates the actual minimum cost to travel from node (i, j) to the destination, and is therefore admissible.

The question remains, however, does the cost function used to generate the route accurately estimate the cost (in terms of crossing time) of traveling along a given route? This question is not easily answered, and it is where historical maritime traffic data can play a key role. As many ships are equipped with an Automatic Identification System (AIS) transponder, which broadcasts the position, heading, and speed of the ship among other information. AIS data is received not only by other ships in the vicinity of a broadcasting ship but also receivers designed to monitor all the maritime traffic in a given area. This data is then made available by various commercial providers, such as AIS Live, or by governmental authorities.

Fig. 4 below compares the generated optimal route with historical AIS routes under the same ice conditions (i.e. same day) as used to generate the route. The sea area shown is primarily the Gulf of Finland with the Finnish coast on the top half of the images and the Estonian coast on the bottom. The AIS routes are shown in yellow, and the generated route in blue. The icebreaker waypoints are shown with a blue star-shaped marker. The calculated transit speeds (i.e. Eq. 11) are indicated by the background color of the sea areas, where the highest speeds are shown in magenta and the lowest speeds shown in cyan. It is apparent that there is significant difference between the historical routes and the generated routes. At this phase in our research, it is not possible for us to say which one of the two routes for each case is more optimal. We can only make a few conjectures and identify aspects requiring further development.

In both cases presented in Fig. 4 the ship represented by the AIS data took a more direct route to the first icebreaker waypoint but spent more time under heavy ice conditions. One possibility is that the crew was not aware of area of open water near the Estonian coast. Another possibility is that they simply believed that the more direct route to the first waypoint would be faster. It is also possibility is that the more direct route actually was faster, compared to the generated route. If this were the case, then it would follow that our cost function is inaccurate. The only way we could determine whether this is the case, however, would be if there was additional AIS data for a similar ship over the same time period, where the ship took this more southerly route. Unfortunately, we have not yet identified a suitable set of AIS data to disambiguate these various possibilities.

One deficiency of the currently implemented method is that it places no constraints on the number and severity of course changes in the calculated route. Ship crews generally would prefer fewer course changes, especially if there is no significant advantage to making such changes. In addition, ships have physical constraints on the rate of course change. It has been suggested that a penalty could be added to the cost function for course changes, the value of which would be proportional to the magnitude of the course change. This would result in straighter routes to be generated, balancing the benefits of changing course with the operational and physical constraints they impose. This approach has not yet been implemented, but it is planned in future work.

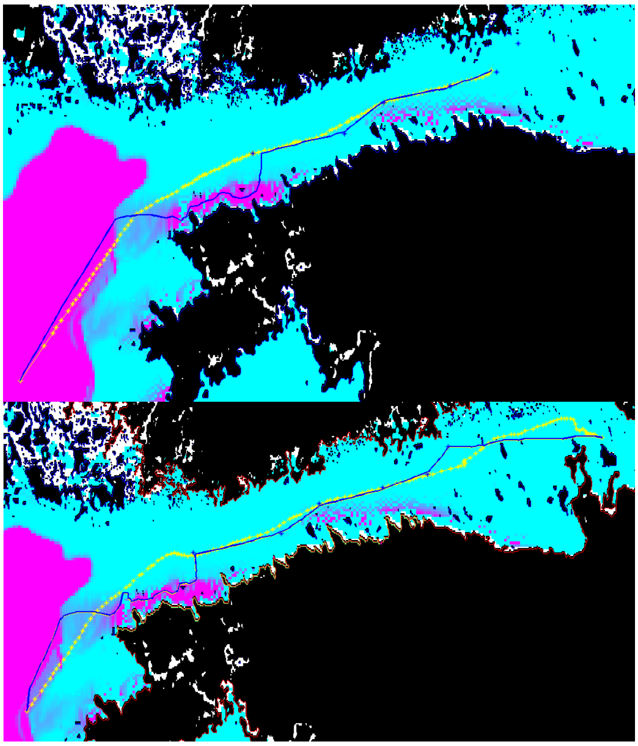

Fig. 4. Comparison of generated routes with historical routes under the same ice conditions. Historical routes, based on AIS data, are shown in yellow with generated routes shown in blue. The cyan and magenta coloring represent transit speed with magenta indicating speeds near those of open water sailing.

VII. Conclusions

This paper has demonstrated the feasibility of ice-aware maritime route optimization. That is, it appears that existing route optimization algorithms, namely the A* algorithm, can be used to minimize the costs associated with sailing through ice-covered waters. Furthermore, the effect of icebreaker assistance can be taken into account by adjusting the cost associated with sailing in the areas between a set of waypoints published by the icebreaker service.

Further work is needed to validate that the results produced using the presented method are indeed optimal. Validation efforts should include further analysis of historical data, as well as simulator-based studies and actual testing of routes at sea. The goal of minimizing crossing time is just one aspect among several that must be taken into account in maritime route planning with others including operational efficiency, safety, and reliability. If the ice information or the cost function used to generate the route are faulty, then the algorithm is prone if not likely to generate suboptimal routes. Thus, it is clear that results can be improved through further work to improve the quality of ice forecasts and further development of models of ship performance in ice. For example, the HELMI sea-ice model provides greater detail than is used in the current model of ship transit speed. Future work should focus on incorporating additional factors, such as ice compression, in order to more accurately predict ship speed under compressive ice conditions. Lastly, future work should also compare the results generated with the presented method to results produced using other algorithms, such as Powell's method.

VIII. Appendix

The following table shows the complete set of neighbors $\mathcal{N}$ and direction vectors $\mathcal{V}$ previously described in Section IIIB. The coordinates reference the relative position from the central node, e.g. (1,-1) is one node to the east and one node to the south from the central node. See Fig. 2 for an illustration.

TABLE III. Set of neighbors and direction vectors employed

#	Coordinates (x,y)	Compass Angle (deg)	#	Coordinates (x,y)	Compass Angle (deg)
1	(0,1)	0	29	(3,2)	56.310
2	(0,-1)	180	30	(3,-2)	123.690
3	(1,0)	90	31	(-3,2)	303.690
4	(-1,0)	270	32	(-3,-2)	236.310
5	(1,1)	45	33	(4,5)	38.6598
6	(1,-1)	135	34	(4,-5)	141.340
7	(-1,-1)	225	35	(-4,5)	321.340
8	(-1,1)	315	36	(-4,-5)	218.660
9	(1,2)	26.565	37	(5,4)	51.340
10	(1,-2)	153.435	38	(5,-4)	128.660
11	(-1,2)	333.435	39	(-5,4)	308.660
12	(-1-,2)	206.565	40	(-5,-4)	231.340
13	(2,1)	63.435	41	(5,1)	78.690
14	(2,-1)	116.565	42	(5,-1)	101.310
15	(-2,1)	296.565	43	(-5,1)	281.310
16	(-2,-1)	243.4349	44	(-5,-1)	258.690
17	(3,1)	71.565	45	(1,5)	11.310
18	(3,-1)	108.435	46	(1,-5)	168.690
19	(-3,1)	288.435	47	(-1,5)	348.690
20	(-3,-1)	251.565	48	(-1,-5)	191.310
21	(1,3)	18.435	49	(1,10)	5.711
22	(1,-3)	161.565	50	(1,-10)	174.289
23	(-1,3)	341.565	51	(-1,10)	354.289
24	(-1,-3)	198.435	52	(-1,-10)	185.711
25	(2,3)	33.690	53	(10,1)	84.289
26	(2,-3)	146.310	54	(10,-1)	95.711
27	(-2,3)	326.310	55	(-10,1)	275.711
28	(-2,-3)	213.690	56	(-10,-1)	264.289

References

[1] Bowditch, N., The American Practical Navigator, National Imagery and Mapping Agency, 2002.

[2] U. Nielsen and J. Jensen, "A novel approach for navigational guidance of ships using onboard monitoring systems," Ocean Engineering, Vol. 38, No. 2-3, 2011, pp. 444-455.

[3] Kotovirta, V., Jalonen, R., Axell, L., Riska, K., & Berglund, R., "A system for route optimization in ice-covered waters," Cold Regions Science and Technology, Vol. 55(1), 2009 pp. 52–62.

[4] BBC News, "Dozens of ships freed from Baltic Sea ice", 5 March 2010. Available: http://bbc.in/1ejqwZJ

[5] Baltic Icebreaking Management, Baltic Sea Icebreaking Report 2009-2010, Available: http://bit.ly/1cpegvx.

[6] G. Hanssen and R. James, "Optimum ship routing", Journal of Navigation, Vol. 13, No. 3, July 1960, pp. 253-272.

[7] M. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," Computer Journal, Vol. 7, No. 2, 1964, pp. 155–162.

[8] Haapala, J., Lönnroth, N., & Stössel, A., "A numerical study of open water formation in sea ice," Journal of Geophysical Research: Oceans, Vol. 110, No. C9, 2005.

[9] Lehtiranta, J., Lensu, M., & Haapala, J., Ice model validation on local scale. Helsinki, 2012.

[10] Neagle, J. N., "Ice resistance prediction and motion simulation for ships operating in the continuous mode of icebreaking," Doctor of Philosophy Thesis, The University of Michigan, Ann Arbor, MI, 1980.

[11] Valanto, P., "The resistance of ship in level ice," Transactions of the Society of Naval Architects and Marine Engineers, Vol. 109, 2001, pp. 53–83.

[12] Liu, J., Lau, M., & Williams, M. F. "Numerical implementation and benchmark of ice-hull interaction model for ship manoeuvring simulations", NRC Publications Archive, Vancouver, BC, Canada, 2008, pp. 215–226.

[13] Sawamura, J., Tsuchiya, H., Tachibana, T., & Osawa, N., "Numerical modeling for ship maneuvering in level ice," Proceedings of 20th International Symposium on Ice (IAHR) Lahti, Finland. 2010, pp. 1–12.

[14] Dick, R. A., Prior, A. D., & Peirce, T. H., "Resistance and propulsion in ice using system identification techniques," Transactions of the Society of Naval Architects and Marine Engineers, Vol. 103, 1995, pp. 237–254.

[15] Montewka, J., Kujala, P., Goerlandt, F., Lensu, M., & Haapala, J., "Modelling a ship performance in dynamic ice. Part I - transforming data into information," Scientific Journal of Warsaw University of Technology, Transport, Vol. 95, 2013, pp. 359–368.

[16] Lindqvist, G., "A straightforward method for calculation of ice resistance of ships," Luleå University of Technology, 1989, pp. 722–735.

[17] La Prairie, D., Wilhelmson, M., & Riska, K., "A transit simulation model for ships in Baltic Ice conditions," Memo No. 200, Espoo, Finland: Helsinki University of Technology, 1995.

[18] Riska, K., Wilhelmson, M., Englund, K., & Leiviskä, T., "Performance of merchant vessels in ice in the Baltic," Espoo, Finland: Helsinki University of Technology, 1997.

[19] Juva, M., & Riska, K., "On the power requirement in the Finnish-Swedish ice class rules" Research Report No. 53 of the Winter Navigation Research Board, Espoo, Finland: Helsinki University of Technology, 2002.

[20] Krupina, N., "Results of time-domain stochastic simulation of ice loads on ship hull from broken ice," Increasing the safety of icebound shipping - Vol. 2, Scientific Report No. 302 of the Helsinki University of Technology Ship Laboratory, Espoo, Finland: Helsinki University of Technology, 2007.

[21] Kaups, K., "Modeling of the ship resistance in compressive level ice," Master Thesis, Espoo, Finland: Aalto University, 2011.

[22] Hart, P. E., Nilsson, N. J., and Raphael, B. "A formal basis for the heuristic determination of minimum cost paths," IEEE Transactions on Systems Science and Cybernetics, Vol. 4, No. 2, 1968, pp. 100–107.

[23] Montewka, J., Sinclair H. "Modelling a ship performance in dynamic ice. Part II - transforming information into knowledge," Scientific Journal of Warsaw University of Technology, Transport, Vol. 95, 2013, pp. 369–382.

[24] R. Dechter and J. Pearl. "The Optimality of A*," In: L. Kanal and V. Kumar, eds., Search in Artificial Intelligence. Springer-Verlag, 1988.