

Aplicação de técnicas de reconhecimento de padrões para estimar o comportamento de bovinos em função de dados de posicionamento GPS

Leandro de Jesus ¹
Celso Correia de Souza ²
Pedro Paulo Pires ³
Luiz Fernando Delboni Lomba ¹
Marcia Ferreira Cristaldo ¹

¹Instituto Federal de Mato Grosso do Sul – IFMS
Rua José Tadao Arima, 222
79200-000 – Aquidauana – MS, Brasil
{leandro.jesus, luiz.lomba, marcia.cristaldo}@ifms.edu.br

²Universidade Anhanguera-Uniderp
Rua Alexandre Herculano, 1400
79037-280 - Campo Grande – MS, Brasil
csouza939@gmail.com

³Embrapa Gado de Corte
Av. Rádio Maia, 830
79106-550 - Campo Grande – MS, Brasil
pedropaulo.pires@embrapa.br

Resumo. Detectar rapidamente animais que precisam de atenção extra devido a problemas de saúde ou bem-estar são essenciais. Estas anormalidades podem ser detectadas pelas mudanças no tempo padrão gasto pelo animal em atividades cotidianas como, ficar de pé, deitar-se, comer ou andar. O objetivo desta pesquisa foi analisar o comportamento de bovinos por meio de monitoramento animal utilizando sensores de posicionamento GPS e algoritmos de reconhecimento de padrões. Os animais foram observados por um analista de domínio (veterinário) e seus comportamentos foram anotados por meio de um software que confere precisão e facilita o treinamento dos algoritmos de classificação supervisionados. Os receptores de GPS, de baixo custo, foram montados em colares e instalados em 11 bovinos no pasto, num total de 42 dias incontinentes, enquanto esses animais foram observados à distância, e os seus comportamentos registrados, num total de 97 horas, em que foram anotados 35.211 comportamentos. Para este conjunto de dados foi obtida uma taxa média de sucesso de 86,1% de classificação de quatro atividades analisadas: pastar, andar, ficar deitado e ficar de pé.

Palavras-chave: gado de corte, pecuária de precisão, etograma, monitoramento animal, bem-estar animal.

Abstract. Rapidly spotting animals that need extra attention to health or wellness issues are essential. These abnormalities can be detected by changes in the standard time spent by the animal in daily activities such as standing, lying, eating or walking. The objective of this research was to analyze the behavior of cattle through animal monitoring using GPS positioning sensors and pattern recognition algorithms. The animals were observed by a domain analyst (veterinary) and their behaviors were annotated through software that confers precision and ease to the training of the supervised classification algorithms. The low-cost GPS receivers were mounted on collars and installed in 11 cattle in the pasture, in a total of 42 non-continuous days, while these animals were observed from a distance, and their systems, for a total of 97 hours, were recorded 35,211 behaviors. For this dataset we obtained an average success rate of 86.1% of the classification of four activities analyzed: grazing, walking, lying and standing.

Keywords: beef cattle, livestock precision, etogram, animal monitoring, animal welfare.

1. Introdução

Os sintomas de alterações de estado de saúde e bem-estar de um animal afetam o seu comportamento (Godsk & Kjaergaard, 2011). Mudanças no tempo padrão gasto pelo animal em atividades como ficar de pé, deitar-se, comer ou andar podem indicar anormalidades causadas por infecções, febre e estresse, social ou ambiental.

Estes comportamentos geralmente são consistentes e previsíveis, mas não podem ser medidos em escala devido ao trabalho necessário para monitorar um grande número de animais de forma contínua, em grandes áreas e/ou de difícil acesso, como o Pantanal.

Os comportamentos mais comumente classificados na literatura são pastar, ficar de pé, deitar-se e andar, entretanto, alguns trabalhos abordam também o comportamento ruminar (Dutta *et al.*, 2015; González *et al.*, 2015; Martiskainen *et al.*, 2009). Robert *et al.* (2009) abordaram apenas os comportamentos ficar de pé, deitar-se e andar, enquanto que Moreau *et al.* (2009), os comportamentos andar, ficar parado e pastar. Handcock *et al.* (2009) avaliaram comportamentos mais complexos do gado bovino, como os dados de movimentação, a cobertura de pastagem e as relações sociais entre os animais do rebanho.

Dos principais sensores utilizados para medir movimentação tem-se como destaque a geotecnologia GNSS (*Global Navigation Satellite System* - Sistema Global de Navegação por Satélite), utilizando o GPS. Oliveira (2013) e Godsk & Kjaergaard (2011) utilizaram exclusivamente este dispositivo, enquanto Moreau *et al.* (2009), González *et al.* (2015) e Dutta *et al.* (2015) utilizaram GPS e acelerômetro. Diosdado *et al.* (2015), Huircán *et al.* (2010), Guo *et al.* (2009), Martiskainen *et al.* (2009), Robert *et al.* (2009) e Scheibe & Gromann (2006) também utilizaram sensores inerciais na medição de movimentação animal.

Neste sentido, ferramentas de monitoramento e análise do comportamento animal se fazem necessárias para auxiliar no manejo do rebanho, na detecção de enfermidades, no melhoramento genético, ou na experimentação de outros gêneros bovinos, sob a hipótese de que a adoção desta tecnologia aceleraria a habilidade em perceber a adaptabilidade do animal ao ambiente (Anderson & Cibils, 2013).

O problema de reconhecimento de atividades bovinas foi abordado a partir de uma perspectiva de software e a classificação foi realizada utilizando-se um conjunto de ferramentas de aprendizagem de máquina, neste caso o *Weka Toolkit* (Hall *et al.*, 2009).

O trabalho apresenta um método para a obtenção de altas taxas de sucesso de classificação por meio da otimização do pré-processamento dos dados brutos de posicionamento GPS e a extração de características que servem como entrada para o algoritmo de aprendizagem automática.

2. Objetivo

Analisar o comportamento de bovinos por meio de monitoramento animal utilizando sensores de posicionamento GPS e algoritmos de reconhecimento de padrões.

3. Material e Métodos

Durante os períodos de coletas de dados para os experimentos os animais utilizados foram mantidos apascentados em pastagem cultivada, cerca de 52 hectares, e com suplementação mineral adequada, água à vontade, tratamento sanitário e manejados pelos profissionais da Embrapa – Gado de Corte (Campo Grande / MS). A área de estudo está localizada na região de planalto da bacia do Alto Paraguai (BAP), sendo de considerável relevância, pois os problemas ambientais da região do Pantanal, em parte, decorrem dos processos ocorridos no planalto.

Da coleta dos dados. Cada animal monitorado foi equipado com um colar. O colar é um dispositivo dotado de um sensor de posicionamento GPS, acelerômetro, giroscópio e magnetômetro, todos de 3 eixos; um sensor de radiação solar e um instrumento de medição de temperatura e humidade relativa do ar. Os dados foram armazenados em um cartão de memória para posterior recuperação manual, momento em que também era trocada a bateria do dispositivo.

Martiskainen *et al.* (2009) coletaram dados de 30 animais por um período de 30 dias, enquanto Diosdado *et al.* (2015) monitoraram 6 animais por um período de apenas 36 horas. Godsk e Kjaergaard (2011) monitoraram 14 vacas leiteiras no pasto por um dia. Neste trabalho, foi utilizada uma amostra de 11 animais em 3 anos diferentes num total de 849 horas de monitoramento.

Poucos estudos dispõem de informações detalhadas sobre os diversos comportamentos do gado bovino, dentre os quais pode-se citar Herbel & Nelson (1966), Zemo & Klemmedson (1970), Hall (1989) e, em especial, o etograma criado por Kilgour *et al.* (2012).

Apesar da diversidade de comportamentos identificáveis os comportamentos mais comuns observados foram: pastar, ficar parado, ruminar, andar, beber água e lamber algum tipo de suplemento (Kilgour *et al.*, 2012). Os três primeiros comportamentos ocupam um período entre 90% e 95% do dia do animal. Assim, definimos os comportamentos a serem classificados: pastar, andar, deitar-se e ficar de pé.

O analista de domínio, neste caso um veterinário, observou os animais e anotou qual atividade o gado estava desempenhando e em qual instante isto ocorreu. Em alguns trabalhos (Godsk & Kjaergaard, 2011; Zheng *et al.*, 2008; Handcock *et al.*, 2009; Nadimi *et al.*, 2012), os pesquisadores realizaram esse processo de observação utilizando caneta e papel. Neste trabalho foi desenvolvido um software de anotação do comportamento, denominado *Ápis*, para ser executado em qualquer dispositivo móvel com sensor GPS e que execute o sistema operacional *Android*.

No primeiro conjunto de dados coletados, chamado de DB1, os quatro bovinos utilizados no experimento foram escolhidos aleatoriamente a partir de um rebanho de 38 animais. Um único animal foi utilizado por coleta. A razão para a utilização de um único bovino em cada coleta foi baseada no desafio prático de observar os animais. É importante salientar que a cada experimento, dos quatro realizados, o animal monitorado era escolhido aleatoriamente pela ordem de chegada no mangueiro, sem a preocupação de utilizar os mesmos animais. Os dados foram coletados no final do mês setembro e durante todo o mês de outubro.

As sequências de dados DB1 se resumiram a um total de 199 horas e 20 minutos de dados de posicionamentos GPS. O analista de domínio realizou a identificação de 3.220 atividades enquanto os dados de GPS estavam sendo coletados pelo sensor nos animais. Desta coleção

de dados de identificação de atividades, a atividade pastar foi classificada 1.641 vezes, seguido por deitar-se com 825 observações, andar com 453 observações e, finalmente, ficar de pé, com 300 observações.

O segundo conjunto de dados foi gerado no ano posterior ao primeiro, definido como DB2. Dois animais da raça Caracu foram equipados com os colares e observados em dois períodos diferentes. O primeiro aconteceu entre os dias 14 e 28 de maio de 2015. O segundo período de coleta do conjunto DB2 ocorreu entre os dias 25 de junho e 06 de julho de 2015.

As sequências de dados DB2 se resumiram em cerca de 513 horas de dados de posicionamento GPS. O analista de domínio realizou mais de 70 horas de observação e a identificação de 31.991 comportamentos, sendo 13.111 pastar, 1.328 beber água, 2.678 andar, 3.685 ficar de pé descansando, 2.007 ficar de pé ruminando, 4.230 ficar deitado descansando e, 4.952 ficar deitado ruminando.

O terceiro período de coleta, DB3, foi realizado nos dias 09 a 10 de maio de 2017 e de 22 a 25 de junho de 2017, com 3 animais. Dificuldades na disponibilização de recursos humanos e materiais impossibilitaram a observação dos animais no terceiro período de coleta. As sequências de dados DB1, DB2 e DB3 estão relacionadas na **Tabela 1**.

Tabela 1. Sequências, de dados coletados, em dias e horas, sobre o comportamento do gado.

Banco de dados	Dados de GPS		Dados de observação	
	No. de dias	No. de horas	No. de dias	No. de horas
DB1	11	199	8	26
DB2	25	513	16	71
DB3	06	137		
Total	42	849	24	97

Sobre a extração de atributos e classificação dos comportamentos. As sequências de dados dos experimentos foram formadas por informações de posicionamento GPS, nas quais os animais executavam as seguintes atividades: pastar, deitar-se, andar ou ficar de pé.

A arquitetura do módulo de classificação do comportamento bovino foi composta por três avaliadores com diferentes estruturas de dados de saídas, inspirado pela pesquisa feita por Zheng *et al.* (2008) e Godsk & Kjaergaard (2011). Os pesquisadores Zheng *et al.* (2008) reconheceram automaticamente diferentes modos de transporte humanos como caminhar, andar de bicicleta e dirigir automóvel a partir de dados brutos de GPS. No trabalho conduzido por Godsk & Kjaergaard (2011), o tratamento de dados brutos de GPS foi utilizado para reconhecer os comportamentos primários de vacas. A suposição básica para os dois trabalhos era de que há um sistema causal, mais ou menos constante, relacionado com o tempo que exerce influência sobre os dados no passado e pode continuar a fazê-lo no futuro. Este sistema causal costuma atuar criando padrões não aleatórios que podem ser detectados em um gráfico de série temporal, em especial, mediante a utilização de um algoritmo de reconhecimento de padrões.

Cada um dos três avaliadores recebia e produzia estruturas de dados específicas. Cada avaliador era configurado individualmente de modo a alterar o comportamento da extração de atributos, de modo que os melhores resultados de reconhecimento de atividades eram obtidos.

Assim, descreveu-se os três blocos de análises: (1) Analisador de Movimento (AM), que transforma os dados brutos de posicionamento GPS e determina o movimento que ocorre entre duas medições sucessivas e é representada por uma estrutura de dados chamada de *Movement Data Structure* (MDS) (Quadro 1); (2) Analisador de Segmento (AS), que agrupa os MDS, organizando-os em segmentos de um determinado tamanho. Todas as informações de movimento são processadas, e como resultado de uma ampla variedade de características

extraídas, é representada na estrutura de dados *Segment Data Structure* (SDS) (Quadro 2); (3) Analisador de Atividades (AA), que usa os SDS como entrada para o algoritmo de aprendizado de máquina e representa a atividade classificada em uma estrutura de dados de atividades chamada *Activity Data Structure* (ADS).

Quadro 1. Resumo dos atributos de um *Movement Data Structure* (MDS).

Identificação do atributo	Descrição do atributo	Descrição dos possíveis valores
Tipo do movimento	Uma representação discreta do movimento executado	Direita, esquerda, para trás, para frente, sem movimento
Ângulo	Ângulo relativo ao movimento anterior	Dado em graus
Magnitude	Distância percorrida entre duas medições	Dado em metros
Velocidade	Velocidade estimada de duas medições em cada instante	Dada em metros por segundo
Direção	Direção absoluta do movimento	Dado em graus
Aceleração	Calculada com base na velocidade estimada de duas medições	Dada em metros por segundo ao quadrado [m/(s*s)]

O Analisador de Movimento (AM) pode ser controlado por meio de três parâmetros de entrada: velocidade mínima, tamanho do histórico e angulação limite. Cada um dos três parâmetros de entrada tem uma influência diferente no AM: velocidade mínima - define o limite entre parado e em movimento; angulação limite - define qual é a fronteira entre os tipos de movimentos: para frente, para trás, esquerda, direita ou parado e, finalmente, o tamanho do histórico é o número de movimentos anteriores utilizados no filtro bayesiano para decidir se existe movimento entre dois pontos adjacentes na série temporal.

No trabalho de Godsk & Kjaergaard (2011), a seleção dos valores dos parâmetros de entrada do AM foi baseada na intuição e experiência de observar o comportamento bovino. Assim, a velocidade mínima foi fixada em 0,3 m/s, o comprimento do histórico (número de pontos de GPS) foi fixado em 4 e a angulação limite foi definida em 40 graus pelo mesmo motivo (Godsk & Kjaergaard, 2011).

Neste trabalho, no entanto, foram variados os valores para velocidade mínima, comprimento do histórico e angulação limite afim de, experimentalmente, encontrar valores adequados para os parâmetros de entrada que maximizavam a taxa de acerto do algoritmo de classificação.

O Analisador de Segmento (AS) calcula os valores de quarenta e seis atributos diferentes, representados no *Segment Data Structure* (SDS) - conjunto de MDS's. Como muitas destas características tendem a ser variantes de outras, elas são agrupadas para maior clareza e listadas no Quadro 2. O critério para a criação do SDS é personalizado por meio do parâmetro de estratégia de segmentação. Por esta estratégia, pode-se variar o tamanho do segmento com o propósito de maximizar a taxa de acerto do algoritmo de classificação.

Quadro 2. Resumo das 46 características geradas pelo analisador de segmentos, agrupadas em 13 atributos.

Parâmetro	Atributo
Tipo do movimento	- Distribuição (% para frente, para trás, direita e esquerda) - Taxa de alteração entre movimento e sem movimento - Taxa de alteração entre qualquer tipo de movimento
Direção	- Acumulado de alterações (para frente, para trás, direita e esquerda) - Taxa de alteração (para frente, para trás, direita e esquerda) - Alteração máxima (para frente, para trás, direita e esquerda)
Velocidade	- Maior, menor e média
Aceleração	- Maior e menor - Média e acumulada (ambas positivas e negativas)

	- Alterações entre positiva e negativa
Distância	- Acumulada para movimento e sem movimento - Maior distância de movimento e sem movimento
Tempo	- Acumulado para movendo e sem movimento

O Analisador de Atividades (AA) atribui a observação feita pelo analista de domínio para cada segmento. O segmento possui a data e hora de início e fim, obtidas por meio do satélite GPS. A observação do analista do domínio, realizada no *Ápis*, na qual ele identificou qual era atividade que o animal estava realizando, também teve registrado sua data e hora por meio do satélite GPS. Assim, foi possível sincronizar os segmentos com a observação.

Baseados nos dados do SDS e na classificação, fornecida pelo analista de domínio, foi possível criar um banco de dados de instâncias (exemplos) previamente classificadas. Este banco de dados de exemplos serviu como entrada para as técnicas de aprendizagem de máquina, e implementadas pelo *Weka Toolkit* (HALL *et al.*, 2009). Foi utilizado, mais especificamente, os algoritmos *END*, *ClassificationViaRegression*, *SVM (SMO)*, *RandomForest*, *J48* e *MultilayerPerceptron*. O problema de reconhecimento das atividades dos bovinos foi tratado utilizando-se um conjunto de algoritmos clássicos de aprendizagem de máquina.

4. Resultados e Discussão

Baseado na metodologia experimental de Godsk & Kjaergaard (2011), no primeiro experimento de classificação definiu-se os parâmetros de configuração de entrada no módulo de análise de atividades da seguinte forma: velocidade mínima foi ajustada para 0,3 m/s, tamanho do histórico foi definido como 4, angulação limite foi definido como 40 graus, a estratégia de segmentação foi definida com o tamanho de segmento de 160 segundos e, finalmente, o classificador END foi selecionado como algoritmo de aprendizado de máquina. Ao utilizar esta configuração foi obtida uma taxa de sucesso de 83,1%. Este resultado é inferior ao obtido no trabalho de Godsk e Kjaergaard (2011), que apresentou uma taxa de acerto de 86,2% no seu melhor caso.

Os parâmetros de configuração para o AA utilizando o algoritmo END foram modificados para verificar a possibilidade de aumento da taxa de acerto. A **Tabela 2** apresenta os parâmetros utilizados, para cada iteração do algoritmo, que compreende a extração dos atributos e a classificação. Os parâmetros são alterados de forma a assumir todas as possibilidades possíveis para um arranjo. O classificador END foi utilizado em sua configuração padrão, ou seja, o software cria 10 conjuntos diferentes de treinamento e testes baseado no conjunto universo de exemplos para cada iteração.

Tabela 2. Parâmetros e valores organizados que foram utilizados em arranjo no módulo de reconhecimento de atividades.

Parâmetros	Mínimo	Máximo	Passo
Analisador de movimentos (MA):			
Velocidade mínima [m/s]	0,1	0,5	0,1
Angulação limite [graus]	10	50	10
Tamanho do histórico	1	10	1
Analisador de segmentos (SA):			
Tamanho do segmento [segundos]	30	180	10

A técnica utilizada tem foco na organização dos dados para obtenção de uma alta taxa de acertos. Assim, combinando os parâmetros, foram obtidas 4.000 formas diferentes de organização dos dados de entrada. O algoritmo END foi executado para as 4.000 iterações

diferentes, e a melhor taxa de acerto (acurácia) obtida foi de 84,7%, com tamanho de segmento igual a 180 segundos, para qualquer valor das variáveis: velocidade mínima, angulação limite e tamanho do histórico.

Foram testados outros algoritmos de aprendizagem de máquina fornecidos pelo *Weka Toolkit*. Na **Tabela 3** são apresentados os resultados obtidos desses testes, tendo definidos os parâmetros de entrada como indicado na melhor classificação. O algoritmo *Random Forest* obteve o melhor desempenho em termos de taxa de instâncias classificadas corretamente (**Tabela 3**). Apesar do algoritmo END não possuir a melhor taxa de acertos ele será utilizado para fins de comparação com o trabalho de Godsk & Kjaergaard (2011).

Tabela 3. Comparação da taxa de acerto (acurácia) entre algoritmos de classificação.

Algoritmo de Aprendizagem de máquina	Taxa de acerto na classificação	Tempo decorrido para classificar as instâncias
END	84,7 %	21 segundos
ClassificationViaRegression	80,9 %	43 segundos
SMO (SVM)	62,3 %	21 segundos
RandomForest	86,1 %	44 segundos
J48	82,2 %	46 segundos
MultilayerPerceptron	72,1 %	477 segundos

Ao avaliar os resultados das classificações, combinando os parâmetros de configuração do analisador de atividades, verificou-se que o tamanho do segmento impacta de forma logarítmica na taxa de acertos para classificação de cada conjunto de instâncias (**Figura 1**). Quanto maior o tamanho do segmento maior a variabilidade do conjunto, elevando a taxa de acertos. Isso confirma a hipótese de que com mais dados sobre a movimentação é possível obter um maior entendimento sobre as atividades realizadas, no entanto, a tendência logarítmica indica que esta tese possui um limite superior. Isto é verificado pela linha de tendência do gráfico na Figura 1.

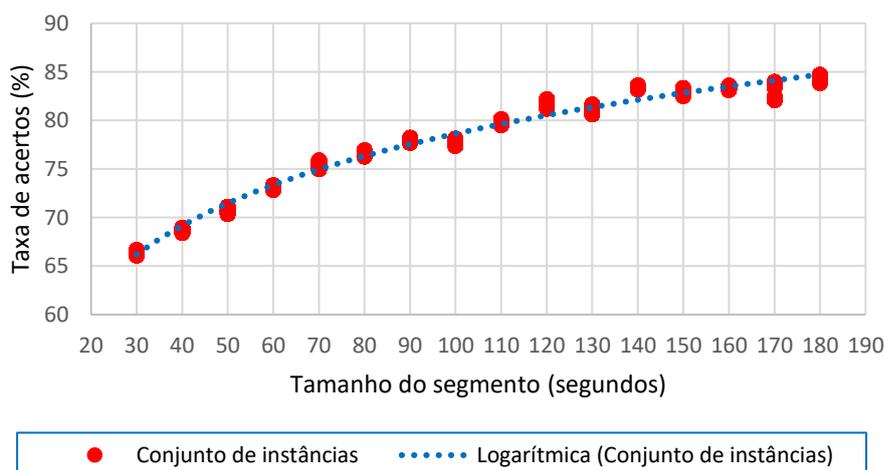


Figura 1. Gráfico da taxa de acertos em função do tamanho do segmento utilizando o algoritmo END. As instâncias (exemplos) classificadas estão agrupadas em função do tamanho do segmento.

No trabalho de Godsk & Kjaergaard (2011), a seleção dos valores dos parâmetros de entrada do AA foi baseada na intuição e experiência de observar o comportamento bovino. No entanto, ao variar os valores das variáveis propostas pelo autor, foi verificado que as

instâncias classificadas estão agrupadas em função do tamanho do segmento. Portanto, as variáveis utilizadas na configuração do pré-processamento, velocidade mínima, angulação limite e tamanho do histórico, não influenciavam significativamente na taxa de acertos, ou ainda, que foi possível dizer que possuíam baixa taxa de correlação com a saída.

Analisando a matriz de confusão (**Tabela 4**) verificou-se, para o melhor caso, que o algoritmo de classificação END concentra seu maior erro na classe que representa a atividade ficar de pé. Lê-se, do conjunto de instâncias que deveriam ser classificadas como ficar de pé, 19 instâncias foram classificadas como pastar, 105 foram classificadas como ficar de pé, 8 como andar e 312 como deitar-se. Percebe-se que o classificador obteve altas taxas de acerto para as atividades andar com 100% e deitar-se com 93,3% de acertos. A atividade ficar de pé contribui para o declínio da taxa média de acertos ao pontuar 70,3%.

Tabela 4. Matriz de confusão da primeira execução utilizando-se todas as classes e o algoritmo END.

Classes	Pastar (a)	Ficar De Pe (b)	Deitar-se (c)	Andar (d)
1 ^a .	321	54	47	23
2 ^a .	14	430	15	2
3 ^a .	19	105	312	8
4 ^a .	0	0	0	527

Analisando a matriz de confusão apresentada por Godsk e Kjaergaard (2011), foi verificado que existe uma grande desproporção entre o número de exemplos de cada classe. Essa situação faz com que os exemplos da classe minoritária sejam classificados incorretamente. Um problema trivial de estatística que deveria ser tratado. Para classes desbalanceadas o autor deveria utilizar outras medidas de avaliação como: sensibilidade (recall), especificidade, precisão, medida F1 ou até mesmo análise ROC. Neste trabalho o problema das classes desbalanceadas foi resolvido utilizando-se as técnicas de reamostragem *SMOTE + Edited Nearest Neighbor Rule (ENN)* (Batista *et al.*, 2004).

Como explanado no artigo de Godsk e Kjaergaard (2011), para classificar uma atividade utilizando dados de GPS, esta atividade por definição deve influenciar no posicionamento do animal, ou seja, a atividade do animal deve apresentar movimento para ser reconhecida. Assim, não se pode ter duas atividades que por definição não apresentam movimento.

A informação de que o animal está realizando o comportamento pastar tem grande relevância na classificação, pois, esta informação serve como base para definir os locais preferenciais para alimentação, ou o tipo de pasto preferido pelo animal, além da possibilidade de identificar os animais que passam mais tempo comendo. Por exemplo, o animal que passa mais tempo pastando e tem pouco ganho de peso, não traz lucro para a propriedade, ou não será um bom candidato em um programa de melhoramento genético. No caso de um animal ficar deitado por muito tempo, supõe-se que está sofrendo de alguma enfermidade. Pode-se inferir este estado quando o animal passa muito tempo sem movimento. Entretanto, ainda não se sabe quando estava de pé ou deitado. Independente do animal estar de pé ou deitado, manter qualquer dos dois estados por muito tempo é motivo para o produtor ou pesquisador monitorar o animal de maneira mais detalhada.

Em outro procedimento experimental procurou-se associar as classes, ficar de pé e deitar-se, em uma única classe, visto que, estas duas classes apresentam como definição a ausência de movimento. Executando o algoritmo END novamente, obteve-se uma melhor taxa de acertos de 88% com os parâmetros com tamanho do segmento igual a 170 segundos. Comparando com a primeira execução, que utilizou todas as classes, obteve-se melhores taxas de acertos do algoritmo END, descritas na **Tabela 5**.

Tabela 5. Melhores taxas de acertos de classificação das atividades do algoritmo END em função do número de classes, quando as atividades deitado e de pé associadas numa mesma classe.

Número de Classes	Taxa de acerto (%)	Tamanho do segmento (s)
3	88,0%	170
4	84,7%	180

Ao avaliar os resultados das classificações, baseadas no algoritmo END, com apenas 3 classes, verificou-se, pela matriz de confusão, **Tabela 6**, que as classes ficar de pé-deitado e a classe andar possuem uma alta taxa de acertos. A classe pastar possui maior erro de classificação porque em certos momentos o animal pastando em movimento e, em alguns momentos, parado.

Tabela 6. Matriz de confusão do algoritmo END a uma taxa de acertos de 88%.

Classes	Pastando (a)	De Pe-Deitado (b)	Andando (c)
1 ^a .	486	127	25
2 ^a .	80	626	5
3 ^a .	1	0	631

Por esta discussão conclui-se que o módulo de análise de atividades comportamentais bovinas pode ser configurado de forma a atender às expectativas específicas de cada pesquisador/produtor. Por exemplo, suponha que se deseja criar um alarme para detectar um animal que está deitado por muito tempo no pasto, pode-se configurar o módulo de análise de atividades de modo que se tenha a melhor taxa de acertos para a classe deitado e, conseqüentemente, obter-se o melhor alarme possível. Assim, o software pode ser refinado para que atenda tais necessidades.

5. Conclusões e Sugestões

Foi apresentado um processo para identificação do comportamento bovino utilizando dados brutos de GPS. Os animais foram observados por um analista de domínio e seus comportamentos foram anotados por meio de um software que confere precisão na anotação do analista de domínio e facilita o treinamento dos algoritmos de classificação supervisionados. Para fornecer dados para o estudo, receptores de GPS de baixo custo foram montados em colares nos bovinos a pasto, num total de 42 dias incontínuos, ou por 849 horas de monitoramento, enquanto eles foram observados a distância, e os seus comportamentos registrados, num total de 97 horas e 35.211 comportamentos anotados.

No processo de reconhecimento de padrões foi apresentado um modelo de pré-processamento dos dados que rendeu uma maior taxa de acertos nos algoritmos de classificação. A técnica consiste na extração de atributos a partir de dados brutos de GPS e análise por meio de algoritmos de reconhecimento padrões.

Dos algoritmos de classificação testados o que obteve o melhor desempenho foi o algoritmo *Random Forest* com 86,1% de acertos contra 84,7% do algoritmo END. Apesar do último apresentar desempenho inferior, seu tempo de treinamento e classificação foi duas vezes menor.

Do ponto de vista do posicionamento GPS, os comportamentos ficar de pé e deitar-se se assemelham, pois são definidos pela ausência de movimento e, portanto, não podem ser diferenciados com eficiência. Assim em um experimento utilizando apenas três classes (pastar, andar e ficar de pé-deitado) foi obtida uma alta taxa de acertos na classificação de 88% com o algoritmo END.

O trabalho apresenta um método eficiente para monitorar e analisar os comportamentos bovinos. A aplicação da metodologia ante a adição de dados de sensores inerciais ao colar solucionará os problemas de similaridade dos comportamentos ficar de pé e ficar deitado, além da possibilidade de reconhecer a atividade de ruminção.

6. Agradecimentos

Agradecemos a Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul - FUNDECT.

7. Referências

- Anderson, R. E. D.; Cibils, A. Spatiotemporal cattle data—a plea for protocol standardization. **Positioning**, v. 4, n. 1, p. 115 – 136, 2013.
- Batista, G.; Prati, R.; Monard, M. A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. **SIGKDD Explorations**, The Association for Computing Machinery (ACM), v. 6, n. 1, p. 20-29, 2004.
- Diosdado, J. V.; Barker, Z.; Hodges, H.; Amory, J.; Croft, D.; Bell, N.; Codling, E. Classification of behaviour in housed dairy cows using an accelerometer based activity monitoring system. **Animal Biotelemetry**, v.3, n.1, p. 15, 2015.
- Dutta, R.; Smith, D.; Rawnsley, R.; Bishop-Hurley, G.; Hills, J.; Timms, G.; Henry, D. Dynamic cattle behavioural classification using supervised ensemble classifiers. **Computers and Electronics in Agriculture**, v. 111, p. 18 – 28, 2015.
- Godsk, T.; Kjaergaard, M. B. High classification rates for continuous cow activity recognition using low-cost gps positioning sensors and standard machine learning techniques. In: Proceedings of the international conference on Advances in data mining: applications and theoretical aspects, 11, 2011. **Anais...** Berlin:Heidelberg, 2011. p. 174–188.
- González, L.; Bishop-Hurley, G.; Handcock, R.; Crossman, C. Behavioral classification of data from collars containing motion sensors in grazing cattle. **Computers and Electronics in Agriculture**, v.110, p.91–102,2015.
- Guo, Y.; Poulton, G.; Corke, P.; Bishop-Hurley, G. J.; Wark, T.; Swain, D. L. Using accelerometer, high sample rate gps and magnetometer data to develop a cattle movement and behaviour model. **Ecological Modelling**, v. 220, n. 17, p. 2068 – 2075, 2009.
- Hall, S. J. Chillingham cattle: social and maintenance behaviour in an ungulate that breeds all year round. **Animal Behaviour**, v. 38, n. 2, p. 215 – 225, 1989.
- Hall, M.; Frank E.; Holmes, G.; Pfahringer, B.; Reutemann, P; Witten, I. H. The weka data mining software: an update. **ACM SIGKDD Explorations Newsletter**, v. 11, n. 1, p. 10–18, 2009.
- Handcock, R. N; Swain, D. L.; Bishop-Hurley, G. J.; Patison, K. P.; Wark, T.; Valencia, P.; Corke, P; O'Neill, C. J. Monitoring animal behaviour and environmental interactions using wireless sensor networks, gps collars and satellite remote sensing. **Sensors**, v. 9, n. 5, p. 3586–3603, 2009.
- Herbel, C. H.; Nelson, A. Activities of hereford and santa gertrudis cattle on a southern new mexico range. **Journal of Range Management**, v. 19, p. 173–176, 1966.
- Huircán, J. I.; Muñoz, C.; Young, H.; Dossow, L. V.; Bustos, J.; Vivallo, G.; Toneatti, M. Zigbee-based wireless sensor network localization for cattle monitoring in grazing fields. **Computers and Electronics in Agriculture**, v. 74, n. 2, p. 258 – 264, 2010.
- Kilgour, R. J.; Uetake, K.; Ishiwata, T.; Melville, G. J. The behaviour of beef cattle at pasture. **Applied Animal Behaviour Science**, v. 138, n. 1–2, p. 12 – 17, 2012.
- Martiskainen, P.; Järvinen, M.; Skön, J.-P.; Tiirikainen, J.; Kolehmainen, M.; Mononen, J. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. **Applied Animal Behaviour Science**, v. 119, n. 1–2, p. 32 – 38, 2009.
- Moreau, M.; Siebert, S.; Buerkert, A.; Schlecht, E. Use of a triaxial accelerometer for automated recording and classification of goats' grazing behaviour. **Applied Animal Behaviour Science**, v.119, n 3–4, p.158– 170, 2009.

Nadimi, E.; Jørgensen, R.; Blanes-Vidal, V.; Christensen, S. Monitoring and classifying animal behavior using zigbee-based mobile ad hoc wireless sensor networks and artificial neural networks. **Computers and Electronics in Agriculture**, v. 82, p. 44 – 54, 2012.

Oliveira, M. T. P. de. **Análise Comportamental de Bovinos Baseada em Trajetórias Semânticas Aplicada à Pecuária de Precisão**. 2013. Campo Grande: UFMS, 2013. 103p. Dissertação Mestrado.

Robert, B.; White, B. J.; Renter, D. G.; Larson, R. L. Evaluation of threedimensional accelerometers to monitor and classify behavior patterns in cattle. **Computers and Electronics in Agriculture**, v.67, n.1-2, p.80–84, 2009.

Scheibe, K.; Gromann, C. Application testing of a new three-dimensional acceleration measuring system with wireless data transfer (WAS) for behavior analysis. **Behavior Research Methods**, v. 38, n. 3, p. 427–433, 2006.

Zemo, T.; Klemmedson, J. Behavior of fistulated steers on a desert grassland. **Journal of Range Management**, Washington, v. 23, n. 3, p. 158–163, 1970.

Zheng, Y.; LIU L.; Wang L.; Xie X. Learning transportation mode from raw gps data for geographic applications on the web. In: Proceedings of the International Conference on World Wide Web, 17, 2008, New York. **Anais...** New York, 2008. p. 247–256.